Probab. Theory Relat. Fields (2014) 158:859–893
DOI 10.1007/s00440-013-0498-8

Mod-discrete expansions

A. D. Barbour · E. Kowalski · A. Nikeghbali

Received: 27 June 2011 / Revised: 24 December 2012 / Published online: 10 April 2013

Abstract In this paper, we consider approximating expansions for the distribution of integer valued random variables, in circumstances in which convergence in law (without normalization) cannot be expected. The setting is one in which the simplest approximation to the n-th random variable X_n is by a particular member R_n of a given family of distributions, whose variance increases with n. The basic assumption is that the ratio of the characteristic function of X_n to that of R_n converges to a limit in a prescribed fashion. Our results cover and extend a number of classical examples in probability, combinatorics and number theory.

Keywords Mod–Poisson convergence · Characteristic function · Poisson–Charlier expansion · Erdős–Kac theorem

Mathematics Subject Classification (2000) 62E17 · 60F05 · 60C05 · 60E10 · 11N60

A. D. Barbour supported in part by Schweizerischer Nationalfonds Projekt Nr. 20-117625/1 and by Australian Research Council Grants Nos. DP120102728 and DP120102398. A. Nikeghbali supported in part by Schweizerischer Nationalfonds Projekt Nr. 200021_119970/1.

A. D. Barbour · A. Nikeghbali (✉)
Institut für Mathematik, Universität Zürich, Winterthurertrasse 190, 8057 Zürich, Switzerland
e-mail: ashkan.nikeghbali@math.uzh.ch

A. D. Barbour
e-mail: a.d.barbour@math.uzh.ch

E. Kowalski
ETH Zurich, D-MATH, Rämistrasse 101, 8092 Zürich, Switzerland
e-mail: kowalski@math.ethz.ch

1 Introduction

The topic of this paper is the explicit approximation, in various metrics, of random variables which, in terms of characteristic functions, behave like a sum

$$X_n = Z_n + Y_n \tag{1.1}$$

of a "model" variable Z_n (for instance, a Poisson random variable) and an independent perturbation Y_n, when the model variable has "large" parameter. Our interest is in discrete random variables, and in cases where this simple-minded decomposition does not in fact exist. We have two motivations:

(1) In probabilistic number theory, it has been known since the proof by Rényi and Turán of the Erdős–Kac theorem that the random variable $\omega(N_n)$ given by the number of prime divisors (without multiplicity, for definiteness) of an integer N_n uniformly chosen in the interval $\{1, 2, \ldots, n\}$ has characteristic function given by

$$\mathbb{E}\{e^{i\theta\omega(N_n)}\} = \mathbb{E}\{e^{i\theta Z_n}\}\Phi(\theta)(1 + o(1))$$

as $n \to \infty$, where $Z_n \sim \text{Po}\,(\log\log n)$ is a Poisson variable with mean $\log\log n$ and $\Phi(\theta)$ is defined by

$$\Phi(\theta) = \frac{1}{\Gamma(e^{i\theta})} \prod_{p \text{ prime}} \left(1 + \frac{e^{i\theta} - 1}{p}\right)\left(1 - \frac{1}{p}\right)^{e^{i\theta}-1},$$

the product being absolutely convergent for all θ real. This $\Phi(\theta)$ is not the characteristic function of a probability distribution, and hence formula (1.1) with $Z_n \sim \text{Po}\,(\log\log n)$ cannot be true. However, we are nonetheless able to obtain explicit approximation statements for the law of $\omega(N_n)$:

Theorem 1.1 *For every integer $r \geq 0$, there exist explicitly computable signed measures $\nu_{r,n}$ on the positive integers such that the total variation distance between the law of $\omega(N_n)$ and $\nu_{r,n}$ is of order $O\{(\log\log n)^{-(r+1)/2}\}$ for $n \geq 3$.*

This is proved in Sect. 7.3, where formulas for the measures $\nu_{1,n}$ and $\nu_{2,n}$ are also given. Such results are new in analytic number theory, where total variation distance estimates have hardly been considered before [but see [4] for a result concerning the total variation distance to a Poisson approximation for the distribution of a truncated version of $\omega(N_n)$].

For more on the significance of the Rényi–Turán formula, comparison with the Keating–Snaith conjectures for the Riemann zeta function, and finite-field analogues, see Kowalski and Nikeghbali [6].

(2) In a beautiful paper, Hwang [5] considered sequences of non-negative integer valued random variables X_n, whose probability generating functions f_{X_n} satisfy

$$e^{\lambda_n(1-z)} f_{X_n}(z) \to g(z),$$

for all $z \in \mathbb{C}$ with $|z| \leq \eta$, for some $\eta > 1$, where the function g is analytic, and $\lim_{n\to\infty} \lambda_n = \infty$. This assumption is also intuitively related to a model (1.1). Under some extra conditions, Hwang exhibits bounds of order $O(\lambda_n^{-1})$ on the accuracy of the approximation of the distribution of X_n by a Poisson distribution with carefully chosen mean, close to λ_n. Hwang [5] also notes that his methods can be applied to families of distributions other than the Poisson family, and gives examples using the Bessel family.

In this paper, we systematically consider sequences of integer valued random variables X_n, whose characteristic functions ϕ_{X_n} satisfy a condition which, in the Poisson context, is some strengthening of the convergence

$$\exp\{\lambda_n(1 - e^{i\theta})\}\phi_{X_n}(\theta) \to \psi(\theta), \quad 0 < |\theta| \leq \pi. \tag{1.2}$$

Under suitable conditions, we derive explicit approximations to the distribution of X_n, in various metrics, by measures related to the Poisson model. The approximations can be made close to any given polynomial order in $\lambda_n^{-1/2}$, if the conditions are sharp enough and the measure is correspondingly chosen. The conditions that we require for these expansions are much weaker than those of Hwang [5]. For instance, his conditions require the X_n to take only non-negative values, and to have exponential tails, neither of which conditions we need to impose.

Our basic result, Proposition 2.1, is very simple and explicit. It enables us to dispense with asymptotic settings, and to prove concrete error bounds. It also allows us to consider approximation by quite general families of distributions on the integers, instead of just the Poisson family, requiring only the replacement of the Poisson characteristic function in (1.2) by the characteristic function corresponding to the family chosen. This enables us to deduce expansions based on any such discrete family of distributions, as shown in Sect. 4, without any extra effort. Indeed, the main problem would seem to be to identify the higher order terms in the expansions, but these turn out simply to be linear combinations of the higher order differences of the basic distribution: see (2.6).

This elementary result, and a simple but powerful theorem that follows from it, are given, together with an example, in Sect. 2. The conditions are then substantially relaxed, in order to allow for wider application, and to treat total variation approximation in a satisfactory manner. The general conclusions are proved in the context of approximating finite signed measures in Sect. 3, and they are reformulated for approximating probability distributions in the usual asymptotic framework in Sect. 4.

In the Poisson context, the measures that result are the Poisson–Charlier measures. Our general results enable us to deduce a Poisson–Charlier approximation with error of order $O(\lambda_n^{-t/2})$, for any prescribed t, assuming that Hwang's conditions hold. We also show that the Poisson–Charlier expansions are valid under more general conditions, in which the X_n may have only a few finite moments. These expansions are established in Sect. 5, and the compound Poisson context is briefly discussed in Sect. 6. We discuss some examples, to sums of independent integer valued random variables, to Hwang's setting and to our first motivation, proving Theorem 1.1, in Sect. 7.

In order to ease the reading of this paper, we give here a diagram indicating the logical dependency of the results we prove. On the left-hand side are the basic approximation theorems, the right-hand side represents applications, and the results of Sect. 4 represent the bridge linking the two:

$$\boxed{\text{Proposition 2.1}}$$

$$\Rightarrow \boxed{\left\{\begin{array}{l}\text{Theorem 2.2}\\ \text{Theorem 3.1}\\ \text{Theorem 3.2}\\ \text{Theorem 3.3}\end{array}\right.} \quad \Rightarrow \boxed{\left\{\begin{array}{l}\text{Corollary 4.1}\\ \text{Proposition 4.2}\end{array}\right.} \quad \Rightarrow \boxed{\left\{\begin{array}{l}\text{Theorem 5.1}\\ \text{Theorem 7.1}\\ \text{Theorem 7.2}\end{array}\right.}$$

We frame our approximations in terms of three distances between (signed) measures μ and ν on the integers: the point metric

$$d_{\mathrm{loc}}(\mu, \nu) \;:=\; \sup_{j\in\mathbb{Z}} |\mu\{j\} - \nu\{j\}|,$$

the Kolmogorov distance

$$d_{\mathrm{K}}(\mu, \nu) \;:=\; \sup_{j\in\mathbb{Z}} |\mu\{(-\infty, j]\} - \nu\{(-\infty, j]\}|,$$

and the total variation norm

$$\|\mu - \nu\| \;:=\; \sum_{j\in\mathbb{Z}} |\mu\{j\} - \nu\{j\}|.$$

Other metrics could also be treated using our methods.

2 The basic estimate

The essence of our argument is the following elementary result, linking the closeness of finite signed measures μ and ν to the closeness of their characteristic functions, when these have a common factor involving a 'large' parameter ρ; for a finite signed measure ζ on $\mathbb{Z}$, the characteristic function ϕ_ζ is defined by $\phi_\zeta(\theta) := \sum_{j\in\mathbb{Z}} e^{ij\theta}\zeta\{j\}$, for $|\theta| \le \pi$.

Proposition 2.1 *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions ϕ_μ and ϕ_ν respectively. Suppose that $\phi_\mu = \psi_\mu\chi$ and $\phi_\nu = \psi_\nu\chi$, and write $d_{\mu\nu} := \psi_\mu - \psi_\nu$. Suppose that, for some $\gamma, \rho, t > 0$,*

$$|d_{\mu\nu}(\theta)| \;\le\; \gamma|\theta|^t \quad \textit{and} \quad |\chi(\theta)| \le e^{-\rho\theta^2} \quad \textit{for all} \quad |\theta| \le \pi. \tag{2.1}$$

Then there are explicit constants α_{1t} and α_{2t} such that

1. $\sup_{j\in\mathbb{Z}} |\mu\{j\} - \nu\{j\}| \le \alpha_{1t}\gamma(\rho \vee 1)^{-(t+1)/2}$;
2. $\sup_{a\le b\in\mathbb{Z}} |\mu\{[a,b]\} - \nu\{[a,b]\}| \le \alpha_{2t}\gamma(\rho \vee 1)^{-t/2}$.

Proof For any $j \in \mathbb{Z}$, the Fourier inversion formula gives

$$\mu\{j\} - \nu\{j\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-ij\theta}(\psi_\mu(\theta) - \psi_\nu(\theta))\chi(\theta)\, d\theta, \tag{2.2}$$

from which our assumptions imply directly that

$$|\mu\{j\} - \nu\{j\}| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \gamma|\theta|^t \exp\{-\rho\theta^2\}\, d\theta.$$

For $\rho \le 1$, we thus have

$$|\mu\{j\} - \nu\{j\}| \le \frac{1}{2\pi}\int_{-\pi}^{\pi} \gamma|\theta|^t\, d\theta \le \frac{\pi^t\gamma}{t+1} =: \beta_{1t}\gamma.$$

For $\rho \ge 1$, it is immediate that

$$|\mu\{j\} - \nu\{j\}| \le \frac{\gamma}{2\pi}\left(\frac{1}{\sqrt{2\rho}}\right)^{t+1}\int_{-\infty}^{\infty} |y|^t e^{-y^2/2}\, dy \le \beta'_{1t}\gamma\rho^{-(t+1)/2},$$

with $\beta'_{1t} := 2^{-(t+1)/2} m_t/\sqrt{2\pi}$; here, m_t denotes the t-th absolute moment of the standard normal distribution. Setting

$$\alpha_{1t} := \max\{\beta_{1t}, \beta'_{1t}\} = \max\left\{2^{-(t+1)/2} m_t/\sqrt{2\pi},\ \pi^t/(t+1)\right\},$$

this proves part 1. The second part is similar, adding (2.2) over $a \le j \le b$, and estimating

$$\frac{\left|e^{-ia\theta} - e^{-i(b+1)\theta}\right|}{|1 - e^{-i\theta}|} \le \frac{\pi}{|\theta|}, \quad |\theta| \le \pi.$$

This gives part 2, with

$$\alpha_{2t} := \max\{2^{-t/2} m_{t-1}\sqrt{\pi/2},\ \pi^t/t\}.$$

□

We shall principally be concerned with taking μ to be the distribution of a random variable X. We allow ν to be a signed measure, because in many cases, such as in the following canonical example and in the Poisson–Charlier expansions of Sect. 5, signed measures appear as the natural approximations.

Let X be an integer valued random variable with characteristic function $\phi_X := \psi\chi$, where χ is the characteristic function of a (well-known) probability distribution R on $\mathbb{Z}$. Suppose that χ satisfies

$$|\chi(\theta)| \leq e^{-\rho\theta^2}, \tag{2.3}$$

as for Proposition 2.1, and that ψ can be approximated by a polynomial expansion around $\theta = 0$ of the form

$$\tilde{\psi}_r(\theta) := \sum_{l=0}^{r} \tilde{a}_l (e^{i\theta} - 1)^l, \tag{2.4}$$

for real coefficients $\tilde{a}_l$ (and with $\tilde{a}_0 = 1$) and some $r \in \mathbb{N}_0$, in that

$$|\psi(\theta) - \tilde{\psi}_r(\theta)| \leq K_{r\delta}|\theta|^{r+\delta}, \quad |\theta| \leq \pi, \tag{2.5}$$

for some $0 < \delta \leq 1$. In view of Proposition 2.1, this suggests that the distribution of X may be well approximated by the signed measure $\nu_r = \nu_r(R; \tilde{a}_1, \ldots, \tilde{a}_r)$ having $\tilde{\psi}_r\chi$ as characteristic function. Now ν_r can immediately be identified as

$$\nu_r = \sum_{l=0}^{r} (-1)^l \tilde{a}_l D^l R, \tag{2.6}$$

where the differences $D^l R$ of the probability measure R are determined by iterating the relation $DR\{j\} := R\{j\} - R\{j-1\}$. Hence, under these assumptions, Proposition 2.1 implies the following theorem; note that the assumption (2.5) is much like supposing that ψ has a Taylor expansion of length r around zero (in powers of $i\theta$), and hence that X has a corresponding number of finite moments.

Theorem 2.2 *Let X be a random variable on $\mathbb{Z}$ with distribution P_X. Suppose that its characteristic function ϕ_X is of the form $\psi\chi$, where χ is the characteristic function of a probability distribution R and satisfies* (2.3) *above. Suppose also that* (2.5) *is satisfied, for some $r \in \mathbb{N}_0$, $\tilde{a}_1, \ldots, \tilde{a}_r \in \mathbb{R}$ and $\delta \geq 0$. Then, writing $t = r + \delta$, we have*

1. $d_{\mathrm{loc}}(P_X, \nu_r) \leq \alpha_{1t} K_{r\delta}(\rho \vee 1)^{-(t+1)/2}$;
2. $d_K(P_X, \nu_r) \leq \alpha_{2t} K_{r\delta}(\rho \vee 1)^{-t/2}$,

with α_{1t} and α_{2t} as in Proposition 2.1, and with $\nu_r = \nu_r(R; \tilde{a}_1, \ldots, \tilde{a}_r)$ as defined in (2.6).

Remark Note that Proposition 2.1 can be applied with $\psi_\mu = 0$, corresponding to μ the zero measure, and $\psi_\nu(\theta) = \tilde{a}_l(e^{i\theta} - 1)^l$, for any $1 \le l \le r$, showing that the contribution from the l-th term in the expansion to $\nu_r\{j\}$ is at most $|\tilde{a}_l|\alpha_{1l}(\rho \vee 1)^{-(l+1)/2}$, and that to $\nu_r\{[a, b]\}$ at most $|\tilde{a}_l|\alpha_{2l}(\rho \vee 1)^{-l/2}$. Thus, if ρ is large and the coefficients $\tilde{a}_l$ moderate, the contributions decrease in powers of $\rho^{-1/2}$ as l increases. In such circumstances, the signed measure ν_r can be seen as a perturbation of the underlying distribution R.

The simplest application of the above results arises when $\phi_X = \phi_Y p_\lambda$, where $p_\lambda(\theta) = e^{\lambda(e^{i\theta}-1)}$ is the characteristic function of the Poisson distribution Po (λ) with mean λ, which satisfies (2.3) with $\rho = 2\pi^{-2}\lambda$, and ϕ_Y is the characteristic function associated with a random variable Y on the integers. In this case, $X = Z + Y$ is the sum of two independent random variables, as in (1.1), with $Z \sim$ Po (λ), and the situation is probabilistically very clear. For $w = w_\theta = e^{i\theta} - 1$, we have $\phi_Y(\theta) = \mathbb{E}\{(1 + w)^Y\}$. The latter expression has an expansion in powers of w up to the term in w^r if the r-th moment of Y exists, with coefficients $\tilde{a}_k := F_k(Y)/k!$, $1 \le k \le r$, where $F_k(Y)$ denotes the k-th factorial moment of Y:

$$F_k(Y) := \sum_{l \ge k} \frac{l!}{(l-k)!}\,\mathbb{P}[Y = l] + \sum_{l \ge 1} (-1)^k \frac{(l+k-1)!}{(l-1)!}\,\mathbb{P}[Y = -l].$$

Thus the asymptotic expansion of X around Po (λ) is simply derived from the factorial moments of the perturbing random variable Y, if they exist.

For example, we could take ϕ_Y to be the characteristic function of a random variable Y_s with distribution

$$\mathbb{P}[Y_s = -l] = s!\,\frac{s}{l(l+1)\ldots(l+s)}, \quad l \ge 1,$$

for some integer $s \ge 1$; the random variable has only $s-1$ moments, and takes negative values, so that the theorems in [5] cannot be applied. However, Y_s has factorial moments

$$F_k(Y_s) = (-1)^k s! \sum_{l \ge 1} \frac{s}{(l+k)\ldots(l+s)} = (-1)^k k!\,\frac{s}{s-k}, \quad 1 \le k \le s-1,$$

and characteristic function

$$\psi_{Y_s}(\theta) = 1 + \sum_{k=1}^{s-1} (-1)^k \frac{s}{s-k}\,(e^{i\theta} - 1)^k - s(1 - e^{i\theta})^s \log(1 - e^{-i\theta}),$$

and (2.5) holds for $\tilde{\psi}_r$ as in (2.4), with $r = s-1$ and any $\delta < 1$, for $\tilde{a}_k = F_k(Y)/k! = (-1)^k s/(s-k)$. Hence, if $X = Z + Y_s$, where $Z \sim$ Po (λ) is independent of Y_s, then Theorem 2.2 can be applied, approximating the distribution of X by the signed measure $\nu_{s-1}(\text{Po}\,(\lambda); \tilde{a}_1, \ldots, \tilde{a}_{s-1})$.

3 Refinements

3.1 Weaker conditions

Proposition 2.1 yields explicit bounds on $d_{\mathrm{loc}}(\mu, \nu)$ and $d_{\mathrm{K}}(\mu, \nu)$ in terms of the quantities specified in (2.1). However, for many applications, a slight weakening of its conditions is useful, in which Conditions (2.1) need not hold either exactly or for all θ, though with corresponding consequences for the bounds obtained. The bound assumed for the difference $\psi_\mu(\theta) - \psi_\nu(\theta)$ in Proposition 2.1 is also replaced by a sum involving different powers of $|\theta|$ in the following theorem. This would at first sight seem superfluous, but is nonetheless useful for asymptotics, when the coefficients of the powers may depend in different ways on the 'large' parameter ρ.

We say that a characteristic function χ is (ρ, θ_0)-*locally normal* if

$$|\chi(\theta)| \le e^{-\rho\theta^2}, \quad 0 \le |\theta| \le \theta_0, \tag{3.1}$$

and that characteristic functions ϕ_μ and ϕ_ν are $(\varepsilon, \eta, \theta_0)$-mod χ *polynomially close*, for some $\varepsilon, \eta > 0$ and $0 < \theta_0 \le \pi$, if $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$, and that, for some $M \ge 0$ and positive pairs $\gamma_m, t_m,\ 1 \le m \le M$,

$$|\psi_\mu(\theta) - \psi_\nu(\theta)| \le \sum_{m=1}^{M} \gamma_m |\theta|^{t_m} + \varepsilon, \quad 0 \le |\theta| \le \theta_0; \tag{3.2}$$

$$|\phi_\mu(\theta) - \phi_\nu(\theta)| \le \eta, \qquad \theta_0 < |\theta| \le \pi. \tag{3.3}$$

Note that, for practical purposes, the quantities ε and η should be as small as possible. Using these definitions, we can state the following theorem, whose proof follows that of Proposition 2.1 very closely, and is omitted.

Theorem 3.1 *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions ϕ_μ and ϕ_ν respectively. Suppose that χ is (ρ, θ_0)-locally normal, and that ϕ_μ and ϕ_ν are $(\varepsilon, \eta, \theta_0)$-mod χ polynomially close. Then, with α_{lt} as for Proposition 2.1, and for any $a_0 < b_0 \in \mathbb{Z}$, we have*

1. $$\sup_{j\in\mathbb{Z}} |\mu\{j\} - \nu\{j\}| \le \sum_{m=1}^{M} \gamma_m \alpha_{1t_m} (\rho \vee 1)^{-(t_m+1)/2} + \tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta;$$

2. $$\sup_{a_0 \le a \le b \le b_0} |\mu\{[a, b]\} - \nu\{[a, b]\}| \le \sum_{m=1}^{M} \gamma_m \alpha_{2t_m} (\rho \vee 1)^{-t_m/2} + (b_0 - a_0 + 1)(\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta),$$

where

$$\tilde{\alpha}_1 := \Big(\frac{\theta_0}{\pi} \wedge \frac{1}{2\sqrt{\pi\rho}}\Big); \quad \tilde{\alpha}_2 := \Big(1 - \frac{\theta_0}{\pi}\Big),$$

and $\gamma_1, \ldots, \gamma_M$ *are as in* (3.2).

The first conclusion yields a bound on $d_{\mathrm{loc}}(\mu, \nu)$. However, the presence of the factor $(b_0 - a_0 + 1)$ in the second bound means that, in contrast to the situation in Proposition 2.1, a direct bound on $d_K(\mu, \nu)$ is not immediately visible. The following result, giving bounds on both $d_K(\mu, \nu)$ and $\|\mu - \nu\|$, is however easily deduced; for a signed measure μ, $|\mu|$ as usual denotes its variation.

Theorem 3.2 *With the notation and conditions of Theorem* 3.1*,*

$$d_{\mathrm{K}}(\mu, \nu) \le \inf_{a \le b} \Big(\varepsilon_{ab}^{(\mathrm{K})} + (|\mu| + |\nu|)\{[a, b]^c\}\Big);$$
$$\|\mu - \nu\| \le \inf_{a \le b} \Big(\varepsilon_{ab}^{(1)} + (|\mu| + |\nu|)\{[a, b]^c\}\Big),$$

where

$$\varepsilon_{ab}^{(\mathrm{K})} := \sum_{m=1}^{M} \gamma_m \alpha_{2t_m} (\rho \vee 1)^{-t_m/2} + (b - a + 1)(\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta);$$
$$\varepsilon_{ab}^{(1)} := (b - a + 1) \left\{ \sum_{m=1}^{M} \gamma_m \alpha_{1t_m} (\rho \vee 1)^{-(t_m+1)/2} + (\tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta) \right\},$$

with α_{lt} *as for Proposition* 2.1 *and with* γ_m *as in* (3.2)*. If also* μ *is a probability measure and* $\nu(\mathbb{Z}) = 1$*, then*

$$d_{\mathrm{K}}(\mu, \nu) \le 2 \inf_{a \le b} \Big(\varepsilon_{ab}^{(\mathrm{K})} + |\nu|\{[a, b]^c\}\Big);$$
$$\|\mu - \nu\| \le \inf_{a \le b} \Big(\varepsilon_{ab}^{(1)} + \varepsilon_{ab}^{(\mathrm{K})} + 2|\nu|\{[a, b]^c\}\Big).$$

Proof The inequality for the total variation norm is immediate. For the Kolmogorov distance, by considering the possible positions of x in relation to $a < b$, we have

$$\begin{aligned}
&|\mu\{(-\infty, x]\} - \nu\{(-\infty, x]\}| \\
&\quad \le \sup_{y<a} |\mu\{(-\infty, y]\} - \nu\{(-\infty, y]\}| + \sup_{a \le y \le b} |\mu\{[a, y]\} - \nu\{[a, y]\}| \\
&\qquad + \sup_{y>b} |\mu\{(b, y]\} - \nu\{(b, y]\}| \\
&\quad \le (|\mu| + |\nu|)\{(-\infty, a) \cup (b, \infty)\} + \varepsilon_{ab}^{(\mathrm{K})}.
\end{aligned}$$

If μ is a probability measure and $\nu(\mathbb{Z}) = 1$, we have

$$|\mu|\{[a,b]^c\} \;=\; 1 - \mu\{[a,b]\} \;\le\; |1 - \nu\{[a,b]\}| + \varepsilon_{ab}^{(\mathrm{K})} \;\le\; |\nu|\{[a,b]^c\} + \varepsilon_{ab}^{(\mathrm{K})}.$$

□

3.2 Sharper total variation approximation

When using Theorem 3.2, it can safely be assumed that the tails of the well-known measure ν can be suitably bounded. However, taking χ to be the characteristic function of the Poisson distribution $\mathrm{Po}\,(\lambda)$, for example, as in the example of Sect. 2, the measure of the tail set $[a,b]^c$ cannot be small unless $b-a$ is large in comparison to $\sqrt{\lambda}$; in an asymptotic sense, as $\lambda \to \infty$ and since $\lambda \asymp \rho$, one would need at least $\rho^{-1/2}(b-a) \to \infty$. As a result, the quantity $\varepsilon_{ab}^{(1)}$ appearing in the bound on the total variation distance would necessarily be of larger asymptotic order than $\sum_{m=1}^M \gamma_m \alpha_{2t_m} \rho^{-t_m/2}$, which, in view of the bound on d_K, would nonetheless seem to be the 'natural' order of approximation. Under somewhat stronger conditions than those of Theorem 3.1, a total variation bound of this order can be deduced (at least, if the quantities ε and η are also suitably small); the argument is reminiscent of that in [8].

We say that a characteristic function χ is $(\rho, \gamma', \theta_0)$-*smoothly locally normal* if $\chi(\theta) := e^{i\zeta\theta - u(\theta)}$ for some $\zeta = \zeta_\chi \in \mathbb{R}$, and for some twice differentiable function u such that $u(0) = u'(0) = 0$, and that

$$|u''(\theta)| \le \gamma'\rho \quad \text{and} \quad \Re\{u(\theta)\} \ge \rho\theta^2, \quad |\theta| \le \theta_0. \tag{3.4}$$

Taking $\chi = p_\lambda$ to be the characteristic function of the Poisson distribution $\mathrm{Po}\,(\lambda)$, for example, we can set $\zeta_\chi = \lambda$ and $u(\theta) = \lambda(1 - e^{i\theta} + i\theta)$, showing that p_λ is (ρ, γ', π)-smoothly locally normal with $\rho = 2\lambda/\pi^2$ and $\gamma' = \pi^2/2$.

For any $\varepsilon, \eta > 0$ and $0 < \theta_0 \le \pi$, we then say that characteristic functions ϕ_μ and ϕ_ν are $(\varepsilon, \eta, \theta_0)$-*smoothly* mod χ *polynomially close* if $\phi_\mu = \psi_\mu \chi$ and $\phi_\nu = \psi_\nu \chi$, and that, for some $M \ge 0$ and positive pairs γ_m, t_m, $1 \le m \le M$, there is a twice differentiable function $\tilde{d}_{\mu\nu}$ defined on $|\theta| \le \theta_0$, for some $0 < \theta_0 \le \pi/4$, such that $\tilde{d}_{\mu\nu}(0) = \tilde{d}'_{\mu\nu}(0) = 0$ and

$$|\tilde{d}''_{\mu\nu}(\theta)| \;\le\; \sum_{m=1}^M \gamma_m |\theta|^{t_m - 2}, \qquad |\theta| \le \theta_0; \tag{3.5}$$

$$e^{-\rho\theta^2} |\psi_\mu(\theta) - \psi_\nu(\theta) - \tilde{d}_{\mu\nu}(\theta)| \;\le\; \varepsilon, \qquad |\theta| \le \theta_0; \tag{3.6}$$

$$|\phi_\mu(\theta) - \phi_\nu(\theta)| \;\le\; \eta, \qquad \theta_0 < |\theta| \le \pi. \tag{3.7}$$

Again, the smaller ε and η, the better the bounds to be obtained.

Theorem 3.3 *Let μ and ν be finite signed measures on $\mathbb{Z}$, with characteristic functions ϕ_μ and ϕ_ν respectively. Suppose that χ is $(\rho, \gamma', \theta_0)$-smoothly locally normal, and* *that* ϕ_μ *and* ϕ_ν *are* $(\varepsilon, \eta, \theta_0)$*-smoothly* mod χ *polynomially close. Assume also that* $\rho \ge 1$ *and that* $\rho\theta_0^2 \ge \log\rho$*. Then there is a function* $\alpha' := \alpha'(t,\gamma)$ *such that*

$$\begin{aligned}\|\mu-\nu\| \;\le\; &\sum_{m=1}^{M}\gamma_m\alpha'(t_m,\gamma')\rho^{-t_m/2} + 3\rho\max\{\varepsilon,\eta\}\\ &+(|\mu|+|\nu|)\{(\lfloor\zeta_\chi\rfloor-\rho,\lfloor\zeta_\chi\rfloor+\rho)^c\},\end{aligned}$$

where γ_m *and* t_m *are as in* (3.5) *and* γ' *is as in* (3.4)*. If* μ *is a probability measure and* $\nu(\mathbb{Z}) = 1$*, then*

$$\|\mu-\nu\| \;\le\; 2\sum_{m=1}^{M}\gamma_m\alpha'(t_m,\gamma')\rho^{-t_m/2}+6\rho\max\{\varepsilon,\eta\}+2|\nu|\{(\lfloor\zeta_\chi\rfloor-\rho,\lfloor\zeta_\chi\rfloor+\rho)^c\}.$$

If (3.5) *and* (3.6) *hold with* $\varepsilon = 0$ *for all* $0 \le |\theta| \le \pi$*, then there is a function* $\alpha^* := \alpha^*(t,\gamma)$ *such that*

$$\|\mu-\nu\| \;\le\; \sum_{m=1}^{M}\gamma_m\alpha^*(t_m,\gamma')\rho^{-t_m/2}.$$

Writing $H := \sum_{m=1}^{M}\gamma_m\rho^{-(t_m+2)/2} + \max\{\varepsilon,\eta\}$, it is clearly enough to show that, for any $j \in (\lfloor\zeta_\chi\rfloor-\rho, \lfloor\zeta_\chi\rfloor+\rho)$,

$$|\mu\{j\}-\nu\{j\}| \;=\; \frac{1}{2\pi}\left|\int_{-\pi}^{\pi} e^{-ij\theta}(\phi_\mu(\theta)-\phi_\nu(\theta))\,d\theta\right| \;\le\; KH, \tag{3.8}$$

for some constant K, giving a total contribution to the bound from such j of order $O(\rho H)$. In view of (3.6) and (3.7), the main effort is to bound $\int_{-\theta_0}^{\theta_0} e^{-\rho\theta^2}|\tilde d_{\mu\nu}(\theta)|d\theta$; however, using (3.5) directly gives a bound of order $O(\rho^{1/2}H)$, which is too large. To get round this, for $|j-\zeta_\chi|$ bigger than $\rho^{1/2}$, we write $e^{-ij\theta}(\phi_\mu(\theta)-\phi_\nu(\theta)) = e^{i(\zeta_\chi-j)\theta-u(\theta)}(\psi_\mu(\theta)-\psi_\nu(\theta))$, and integrate (3.8) twice by parts, to get a factor of $(j-\zeta_\chi)^2$ in the denominator. To make this argument work, we need to continue the function $\tilde w(\theta) := e^{-u(\theta)}\tilde d_{\mu\nu}(\theta)$ into $\theta_0 < |\theta| \le \pi$ in suitable fashion. For this, we use the following technical lemma, whose proof is given in the Appendix.

Lemma 3.4 *Let* $w\colon(-\infty,0]\to\mathbb{R}$ *be such that* $w(0)=a$ *and* $w'(0)=b$*. Then* w *can be continued differentiably on* $[0,\infty)$ *by a piecewise quadratic function such that* $|w''(x)| \le c$ *for all* $x > 0$ *for which* $w''(x)$ *is defined, and such that* $w(x) = 0$ *for all*

$$x \;\ge\; \frac{1}{c}\left\{|b| + 2\sqrt{|ac+\tfrac{1}{2}\mathrm{sgn}(b)b^2|}\right\};$$

furthermore, $\max_{x\ge0}|w(x)| \le |a| + b^2/2c$*.*

We then write

$$\mu\{j\} - \nu\{j\} = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\left\{e^{-u(\theta)}[d_{\mu\nu}(\theta) - \tilde d_{\mu\nu}(\theta)] + \tilde w(\theta)\right\} d\theta, \quad (3.9)$$

where $d_{\mu\nu} := \psi_\mu - \psi_\nu$, and, for each j, bound the two parts of the final expression separately.

Proof of Theorem 3.3 (i). For the first step, we use Lemma 3.4 to continue the real and imaginary parts of $\tilde w(\theta)$ into $\theta_0 \le |\theta| \le \pi$, in such a way that $\tilde w$ is piecewise twice differentiable on $[-\pi, \pi]$ and satisfies

$$\tilde w(-\pi) = \tilde w(\pi) = \tilde w'(-\pi) = \tilde w'(\pi) = 0, \quad (3.10)$$

with the second derivatives of the real and imaginary parts suitably bounded. Since

$$\tilde w'(\theta) = e^{-u(\theta)}\{\tilde d'_{\mu\nu}(\theta) - u'(\theta)\tilde d_{\mu\nu}(\theta)\},$$

it follows from (3.4) and (3.5) that

$$|\tilde w(\theta_0)| \le \sum_{m=1}^{M} \frac{\gamma_m}{t_m(t_m-1)}\theta_0^{t_m} e^{-\rho\theta_0^2} \le \sum_{m=1}^{M} |a_m|; \quad (3.11)$$

$$|\tilde w'(\theta_0)| \le \sum_{m=1}^{M} \frac{\gamma_m}{t_m(t_m-1)}\theta_0^{t_m-1} e^{-\rho\theta_0^2}\{t_m + \gamma'\rho\theta_0^2\} \le \sum_{m=1}^{M} |b_m|, \quad (3.12)$$

where

$$|a_m| := t_m^{-1}\gamma_m\kappa_1(t_m, \gamma')\theta_0^{t_m} e^{-\rho\theta_0^2}, \quad |b_m| := \gamma_m\kappa_1(t_m, \gamma')\rho\theta_0^{t_m+1} e^{-\rho\theta_0^2}, \quad (3.13)$$

and $\kappa_1(t, \gamma) := (t + \gamma)/\{t(t-1)\}$. Hence we can continue $\tilde w$ in $\theta_0 \le \theta \le \pi$ by a sum of functions $\sum_{m=1}^M \tilde w_m$, where $|\tilde w_m(\theta_0)| \le |a_m|$ and $|\tilde w'_m(\theta_0)| \le |b_m|$ for each m, and these bounds at θ_0 hold also for the real and imaginary parts $\tilde w_{mr}$ and $\tilde w_{mi}$ of $\tilde w$. Define $\tilde w_{mr}$ and $\tilde w_{mi}$ in $\theta_0 \le \theta \le \pi$ using Lemma 3.4, in each case restricting their second derivatives by taking

$$c_m := 4\gamma_m\kappa_1(t_m, \gamma')\rho^2\theta_0^{t_m+2} e^{-\rho\theta_0^2}. \quad (3.14)$$

Then it follows from the lemma that the length of the θ-interval beyond θ_0 on which $\tilde w_m$ is not identically zero is bounded by

$$\frac{1}{c_m}\left\{|b_m|(1+\sqrt{2}) + 2\sqrt{|a_m|c_m}\right\} \le \frac{1+3\sqrt{2}}{4\rho\theta_0} \le \ell := \frac{2}{\rho\theta_0}, \quad (3.15)$$

from (3.11) and (3.12), the bound being the same for all m; note that

$$\ell \le \frac{2\theta_0}{\rho\theta_0^2} \le \frac{\pi}{2},$$

since $\theta_0 \le \pi/4$ and $\rho\theta_0^2 \ge 1$. From this and (3.14), and from the analogous continuation in $-\pi \le \theta \le -\theta_0$, it follows also that

$$\int_{\theta_0<|\theta|\le\pi} |\tilde{w}''_m(\theta)|\,d\theta \le 4\ell c_m \le 32\gamma_m\kappa_1(t_m,\gamma')\rho\theta_0^{t_m+1}e^{-\rho\theta_0^2}, \tag{3.16}$$

and, using (3.11), (3.12), (3.13) and Lemma 3.4, that

$$\rho\int_{\theta_0<|\theta|\le\pi} |\tilde{w}_m(\theta)|\,d\theta \le 4\ell\rho\{|a_m| + b_m^2/2c_m\} \le 5\,\gamma_m\kappa_1(t_m,\gamma')\theta_0^{t_m-1}e^{-\rho\theta_0^2}. \tag{3.17}$$

(ii). The next step is to bound the first part of the integral in (3.9). Here, by (3.6) and (3.7), we have $|e^{-u(\theta)}[d_{\mu\nu}(\theta) - \tilde{d}_{\mu\nu}(\theta)]| \le \varepsilon$ in $|\theta| \le \theta_0$, whereas, in $\theta_0 < |\theta| \le \pi$, it is bounded by $\eta + |\tilde{w}(\theta)|$. Hence, for any j, we use (3.17) to give

$$\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}e^{-u(\theta)}[d_{\mu\nu}(\theta) - \tilde{d}_{\mu\nu}(\theta)]\,d\theta\right| \le \max\{\varepsilon,\eta\} + \frac{5}{2\pi\rho}\sum_{m=1}^{M}\gamma_m\kappa_1(t_m,\gamma')\theta_0^{t_m-1}e^{-\rho\theta_0^2}. \tag{3.18}$$

Noting also that, if $\rho\theta^2 \ge \log\rho \ge 0$, $\theta > 0$ and $t \ge 2$, then

$$\rho^{t/2}\theta^{t-1}e^{-\rho\theta^2} = \{\rho e^{-\rho\theta^2/2}\}^{1/2}(\rho\theta^2)^{(t-1)/2}e^{-\rho\theta^2/2} \le k_2(t),$$

for $k_2(t) = \{(t-1)/e\}^{(t-1)/2}$, it follows that

$$\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}e^{-u(\theta)}[d_{\mu\nu}(\theta) - \tilde{d}_{\mu\nu}(\theta)]\,d\theta\right| \le \max\{\varepsilon,\eta\} + \frac{5}{2\pi\rho}\sum_{m=1}^{M}\gamma_m\kappa_1(t_m,\gamma')k_2(t_m)\rho^{-t_m/2}. \tag{3.19}$$

This bounds the first element of (3.9) as $O(H)$ for all j.

(iii). For the second part of (3.9), we begin by considering values of j such that $|j-\zeta_\chi| < 1+\lceil\sqrt{\rho}\rceil$. Here, we write

$$
\begin{aligned}
\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\,\tilde{w}(\theta)\,d\theta\right| \;&\le\; \frac{1}{2\pi}\int_{-\pi}^{\pi} |\tilde{w}(\theta)|\,d\theta \\
&\le \frac{1}{2\pi}\int_{|\theta|\le\theta_0} e^{-\rho\theta^2}|\tilde{d}_{\mu\nu}(\theta)|\,d\theta + \frac{1}{2\pi}\sum_{m=1}^{M}\int_{\theta_0<|\theta|\le\pi} |\tilde{w}_m(\theta)|\,d\theta.
\end{aligned}
$$

Since, by (3.5),

$$
|\tilde{d}'_{\mu\nu}(\theta)| \;\le\; \sum_{m=1}^{M}\frac{\gamma_m}{t_m-1}|\theta|^{t_m-1} \quad\text{and}\quad |\tilde{d}_{\mu\nu}(\theta)| \;\le\; \sum_{m=1}^{M}\frac{\gamma_m}{t_m(t_m-1)}|\theta|^{t_m}, \tag{3.20}
$$

the first integral is bounded, as in the proof of Proposition 2.1, by

$$
\sum_{m=1}^{M}\gamma_m\frac{\alpha_{1t_m}}{t_m(t_m-1)}\,\rho^{-(t_m+1)/2}, \tag{3.21}
$$

and the second is bounded, as above, by

$$
\frac{5}{2\pi\rho}\sum_{m=1}^{M}\gamma_m\kappa_1(t_m,\gamma')k_2(t_m)\rho^{-t_m/2},
$$

giving the bound

$$
\begin{aligned}
&\left|\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\,\tilde{w}(\theta)\,d\theta\right| \\
&\qquad\le \sum_{m=1}^{M}\gamma_m\left\{\frac{\alpha_{1t_m}}{t_m(t_m-1)} + \frac{5}{2\pi}\kappa_1(t_m,\gamma')k_2(t_m)\right\}\rho^{-(t_m+1)/2},
\end{aligned} \tag{3.22}
$$

since also $\rho \ge 1$. The bound is of order $O(\rho^{1/2}H)$, but there are only at most $4+2\sqrt{\rho} \le 6\sqrt{\rho}$ integers j satisfying $|j-\zeta_\chi| < 1+\lceil\sqrt{\rho}\rceil$, so that their sum is of order $O(\rho H)$, which is as required.

(iv). For $|j-\zeta_\chi| \ge 1+\lceil\sqrt{\rho}\rceil$, integrating twice by parts and using (3.10), it follows that

$$
\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\,\tilde{w}(\theta)\,d\theta \;=\; -\frac{1}{2\pi(j-\zeta_\chi)^2}\int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\,\tilde{w}''(\theta)\,d\theta, \tag{3.23}
$$

where

$$\tilde{w}''(\theta) \;=\; \left(\tilde{d}''_{\mu\nu}(\theta) - 2\tilde{d}'_{\mu\nu}(\theta)u'(\theta) + \tilde{d}_{\mu\nu}(\theta)\{(u'(\theta))^2 - u''(\theta)\}\right)\, e^{-u(\theta)} \tag{3.24}$$

in $|\theta| \le \theta_0$. Hence, using (3.5), (3.20) and the fact that, from (3.4), $|u'(\theta)| \le \gamma'\rho|\theta|$ in $|\theta| \le \theta_0$, the part of the integral in (3.23) for this range of θ can be bounded by

$$\begin{aligned}
&\left| \int\limits_{\theta \le \theta_0} e^{-i(j-\zeta_\chi)\theta}\, \tilde{w}''(\theta)\, d\theta \right| \\
&\le \sum_{m=1}^{M} \int\limits_{\theta \le \theta_0} \gamma_m \left\{ |\theta|^{t_m-2} + \frac{2\gamma'\rho}{t_m-1}\, |\theta|^{t_m} + \frac{\gamma'\rho}{t_m(t_m-1)}\, |\theta|^{t_m}(1+\gamma'\rho\theta^2) \right\} e^{-\rho\theta^2}\, d\theta \\
&\le \sum_{m=1}^{M} \gamma_m\, \beta'(t_m,\gamma')\rho^{-(t_m-1)/2},
\end{aligned} \tag{3.25}$$

after some calculation, where, with m_t as in Proposition 2.1,

$$\beta'(t,\gamma') \;:=\; \frac{m_{t-2}}{4t\, 2^{t/2}\sqrt{\pi}}\{4t + 2(2t+1)\gamma' + (t+1)(\gamma')^2\}.$$

The remaining part of the integral in (3.23), for $\theta_0 < |\theta| \le \pi$, yields an additional element of

$$\begin{aligned}
\sum_{m=1}^{M} \int\limits_{\theta_0 < |\theta| \le \pi} |\tilde{w}''_m(\theta)|\, d\theta &\le 32 \sum_{m=1}^{M} \gamma_m \kappa_1(t_m,\gamma')\rho\theta_0^{t_m+1} e^{-\rho\theta_0^2} \\
&\le 32 \sum_{m=1}^{M} \gamma_m \kappa_1(t_m,\gamma')\, k_3(t_m)\rho^{-(t_m-1)/2},
\end{aligned} \tag{3.26}$$

from (3.16), with $k_3(t) := \{(t+1)/2e\}^{(t+1)/2}$. As a result, we find that, for $|j - \zeta_\chi| \ge 1 + \lceil \sqrt{\rho} \rceil$, the second part of (3.9) can be bounded by

$$\begin{aligned}
&\left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta}\, \tilde{w}(\theta)\, d\theta \right| \\
&\quad \le \frac{1}{(j-\zeta_\chi)^2} \sum_{m=1}^{M} \gamma_m \left\{ \beta'(t_m,\gamma') + \frac{16}{\pi}\kappa_1(t_m,\gamma')\, k_3(t_m) \right\} \rho^{-(t_m-1)/2},
\end{aligned} \tag{3.27}$$

and adding over $|j - \zeta_\chi| \ge 1 + \lceil \sqrt{\rho} \rceil$ gives a contribution of order $O(\rho H)$.

(v). The final step is to make the arbitrary choice $s = \rho$ in the bound

$$\|\mu - \nu\| \le \sum_{|j-\lfloor\zeta_\chi\rfloor|<s} |\mu\{j\} - \nu\{j\}| + (|\mu| + |\nu|)\{(\lfloor\zeta_\chi\rfloor - s, \lfloor\zeta_\chi\rfloor + s)^c\},$$

and to note that, if μ is a probability measure and $\nu(\mathbb{Z}) = 1$, then

$$\begin{aligned} |\mu|\{(a,b)^c\} = 1 - \mu\{(a,b)\} &\le |1 - \nu\{(a,b)\}| + |\nu\{(a,b)\} - \mu\{(a,b)\}| \\ &\le |\nu|\{(a,b)^c\} + \sum_{a<j<b} |\mu\{j\} - \nu\{j\}|. \end{aligned}$$

(vi). If (3.5) and (3.6) hold with $\varepsilon = 0$ for all $0 \le |\theta| \le \pi$ (implying, in particular, that η is irrelevant), the proof simplifies dramatically. The considerations concerning $\theta_0 < |\theta| \le \pi$ become unnecessary. This leaves the bound

$$|\mu(j) - \nu(j)| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-i(j-\zeta_\chi)\theta} \tilde{w}(\theta)\, d\theta \right| \le \sum_{m=1}^{M} \frac{\gamma_m \alpha_{1t_m}}{t_m(t_m-1)} \rho^{-(t_m+1)/2} \tag{3.28}$$

for $|j - \zeta_\chi| < 1 + \lceil\sqrt{\rho}\rceil$, where $\tilde{w}(\theta) = e^{-u(\theta)} d_{\mu\nu}(\theta)$. Then, since $e^{-i(j-\zeta_\chi)\theta}\tilde{w}(\theta)$ is a 2π-periodic function, the integration by parts in (3.23) remains true, giving the bound

$$|\mu(j) - \nu(j)| \le \frac{1}{(j-\zeta_\chi)^2} \sum_{m=1}^{M} \gamma_m \beta'(t_m, \gamma') \rho^{-(t_m-1)/2} \tag{3.29}$$

for $|j - \zeta_\chi| \ge 1 + \lceil\sqrt{\rho}\rceil$. Adding over all j gives the final bound, with $\alpha^*(t, \gamma') := 2\beta'(t, \gamma') + 6\alpha_{1t}/\{t(t-1)\}$. □

In certain applications, the difference $d_{\mu\nu}$ is expressed in the form $d_{\mu\nu}(\theta) = \hat{d}_{\mu\nu}(e^{i\theta} - 1)$. If it is true that $\hat{d}_{\mu\nu}(0) = \hat{d}'_{\mu\nu}(0) = 0$ and $|\hat{d}''_{\mu\nu}(w)| \le \hat{\gamma}|w|^{t-2}$ for complex w such that $|w| \le \theta_0$, then it follows that $d_{\mu\nu}(0) = d'_{\mu\nu}(0) = 0$ and that

$$|d''_{\mu\nu}(\theta)| \le \left(1 + \frac{2 \wedge \theta_0}{t-1}\right) \hat{\gamma}|\theta|^{t-2}, \qquad |\theta| \le \theta_0. \tag{3.30}$$

4 Approximating probability distributions

4.1 The general case

The most common application of the general bounds is when μ is a probability distribution which is close to a member R_λ of a family $\{R_\lambda, \lambda > 0\}$ of probability distributions on the integers, and one is interested in bounds when λ is large. Suppose, in particular, that the characteristic function r_λ of R_λ is (ρ, γ', π)-smoothly locally normal, and that $\phi_\mu = \psi r_\lambda$, where ψ has a polynomial approximation $\tilde{\psi}_r$ as given in (2.4), for some $r \in \mathbb{N}$ and $\tilde{a}_1, \ldots, \tilde{a}_r \in \mathbb{R}$. This indicates that μ may be close to $\nu = \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ given in (2.6). The following corollary, in which we use a more probabilistic notation for μ, establishes the corresponding results.

Corollary 4.1 *Let X be an integer valued random variable with distribution P_X and characteristic function $\phi_X := \psi r_\lambda$, where r_λ is a $(\rho, \gamma', \theta_0)$-smoothly locally normal characteristic function and $\rho \geq 1$. Let $\tilde{\psi}_r$ be as in (2.4). Then, if ϕ_X and $\tilde{\psi}_r r_\lambda$ are $(\varepsilon, \eta, \theta_0)$-*mod *$r_\lambda$ polynomially close, it follows that*

1. $d_{\mathrm{loc}}(P_X, \nu_r) \leq \sum_{m=1}^{M} \gamma_m \alpha_{1t_m} \rho^{-(t_m+1)/2} + \tilde{\alpha}_1 \varepsilon + \tilde{\alpha}_2 \eta$;
2. $d_K(P_X, \nu_r) \leq 2 \inf_{a \leq b} \left(\varepsilon_{ab}^{(\mathrm{K})} + |\nu_r|\{[a,b]^c\} \right)$;
3. $\|P_X - \nu_r\| \leq \inf_{a \leq b} \left(\varepsilon_{ab}^{(1)} + \varepsilon_{ab}^{(\mathrm{K})} + 2|\nu_r|\{[a,b]^c\} \right)$,

where the quantities appearing in the bounds are as in Theorem 3.2, and with $\nu_r = \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ as defined in (2.6). Furthermore, if ϕ_X and $\tilde{\psi}_r r_\lambda$ are $(\varepsilon, \eta, \theta_0)$-smoothly mod *$r_\lambda$ polynomially close, then*

4. $$\|P_X - \nu_r\| \leq 2 \sum_{m=1}^{M} \gamma_m \alpha'(t_m, \gamma') \rho^{-t_m/2} + 6\rho \max\{\varepsilon, \eta\} + 2|\nu_r|\{(\lfloor \zeta_{r_\lambda} \rfloor - \rho, \lfloor \zeta_{r_\lambda} \rfloor + \rho)^c\},$$

and, if (3.5) *and* (3.6) *hold with $\varepsilon = 0$ for all $0 \leq |\theta| \leq \pi$, then*

5. $\|P_X - \nu_r\| \leq \sum_{m=1}^{M} \gamma_m \alpha^*(t_m, \gamma') \rho^{-t_m/2}$.

Remark Taking $\psi_\mu = 0$ and $\psi_\nu = (e^{i\theta} - 1)^l$ in Theorem 3.3 for $l \geq 2$ gives $|d''_{\mu\nu}(\theta)| \leq l(l+1)|\theta|^{l-2}$ for all θ and $d'_{\mu\nu}(0) = 0$, where $d_{\mu\nu}(\theta) = \psi_m(\theta) - \psi_n(\theta)$. Hence, by the final part of the theorem, the contribution from the l-th term in the signed measure ν_r of (2.6) has total variation norm at most $\alpha^*(l, \gamma') l(l+1)|\tilde{a}_l|\rho^{-l/2}$, for $2 \leq l \leq r$.

4.2 Probability distributions as approximations

The use of signed measures to approximate probability distributions is convenient, but not very natural. However, the signed measures $\nu_1(R_\lambda; \tilde{a}_1)$ and $\nu_2(R_\lambda; \tilde{a}_1, \tilde{a}_2)$ can often be replaced by suitably translated members of the family $\{R_\lambda,\ \lambda > 0\}$, with the same *asymptotic* rate of approximation, by fitting the first two moments, a procedure analogous to that used in the Berry–Esseen theorem. We accomplish this under some further mild assumptions on the distributions R_λ.

We call the family $\{R_\lambda,\ \lambda > 0\}$ *amenable* if the following three conditions are satisfied. First, the characteristic functions r_λ are to be $(\rho(\lambda), \gamma', \pi)$-smoothly locally normal (with the same value of γ' for all), where $\lim_{\lambda\to\infty} \rho(\lambda) = \infty$; secondly, if

$b_1 := b_1(\lambda, \lambda')$ and $b_2 := b_2(\lambda, \lambda')$ are chosen to make the first two derivatives of the function

$$w_{\lambda,\lambda'}(\theta) \;:=\; r_{\lambda'}(\theta) - r_\lambda(\theta)\{1 + b_1(e^{i\theta} - 1) + b_2(e^{i\theta} - 1)^2\} \tag{4.1}$$

vanish at zero ($w_{\lambda,\lambda'}(\theta) = 0$ is automatic), then $\delta_{\lambda,\lambda'}(\theta) := w_{\lambda,\lambda'}(\theta)/r_\lambda(\theta)$ is to satisfy

$$|\delta''_{\lambda,\lambda'}(\theta)| \;\le\; f(|\lambda - \lambda'|)|\theta|, \quad |\theta| \le \pi, \tag{4.2}$$

for some continuous function $f\colon \mathbb{R}_+ \to \mathbb{R}_+$; and thirdly, if $Z_\lambda \sim R_\lambda$, then $\mu(\lambda) := \mathbb{E}Z_\lambda$ and $\sigma^2(\lambda) := \mathrm{Var}\, Z_\lambda$ should exist, with $\sigma^2(\cdot)$ strictly increasing from zero to infinity, and the functions $\mu(\cdot)$, $\sigma^2(\cdot)$ and $(\sigma^2)^{-1}(\cdot)$ are all to be uniformly continuous.

The quantities b_1 and b_2, as defined in (4.1), can be explicitly expressed:

$$b_1(\lambda, \lambda') \;=\; \mu(\lambda') - \mu(\lambda); \quad b_2(\lambda, \lambda') \;=\; \tfrac{1}{2}\{\sigma^2(\lambda') - \sigma^2(\lambda) - b_1(1 - b_1)\}, \tag{4.3}$$

and it follows from (4.2) that

$$|\delta_{\lambda,\lambda'}(\theta)| \;\le\; \tfrac{1}{6} f(|\lambda - \lambda'|)|\theta|^3, \quad |\theta| \le \pi. \tag{4.4}$$

Note that the Poisson family $\{\mathrm{Po}\,(\lambda)\; \lambda > 0\}$ is amenable.

Now the signed measures ν_r, $r \ge 2$, have mean and variance given by

$$\mu_* = \mu(\lambda) + \tilde{a}_1; \quad \sigma_*^2 \;=\; \sigma^2(\lambda) + 2\tilde{a}_2 + \tilde{a}_1(1 - \tilde{a}_1), \tag{4.5}$$

and the corresponding equations for ν_1 just have $\tilde{a}_2 = 0$. However, when choosing a translation of R_λ to match these moments, only integer translations m of R_λ can be allowed, since the distributions must remain on the integers, and so it is not possible to match both moments exactly within the family. To circumvent this, we extend to approximation by a member of the family of probability distributions $Q_{mp}(R_{\lambda'})$, for $\lambda' > 0$, $m \in \mathbb{Z}$ and $0 \le p < 1$, where

$$Q_{mp}(R_{\lambda'})\{j\} \;:=\; pR_{\lambda'}\{j - m - 1\} + (1 - p)R_{\lambda'}\{j - m\}. \tag{4.6}$$

If $Z \sim R_{\lambda'}$, then $Q_{mp}(R_{\lambda'})$ is the distribution of $Z + m + I$, where $I \sim \mathrm{Be}\,(p)$ is independent of Z. $Q_{mp}(R_{\lambda'})$ has characteristic function $q_{mp}(R_{\lambda'})$ given by

$$q_{mp}(R_{\lambda'})(\theta) \;:=\; e^{im\theta}(1 + p(e^{i\theta} - 1))r_{\lambda'}(\theta), \tag{4.7}$$

similar to the measure $\nu_2\{R_{\lambda'}; m + p, \binom{m}{2} + mp\}$, but with terms of higher order as powers of $(e^{i\theta} - 1)$ as well.

Among the distributions $\{Q_{mp}(R_{\lambda'});\ \lambda' > 0, m \in \mathbb{Z}, 0 \le p < 1\}$, we can always find one having a given mean μ_* and variance σ_*^2, provided that $\{R_\lambda,\ \lambda > 0\}$ is amenable and that $\sigma_*^2 \ge 1/4$, by solving the equations

$$\mu_* = \mu(\lambda') + m + p; \quad \sigma_*^2 \ = \ \sigma^2(\lambda') + p(1-p). \tag{4.8}$$

To do so, let λ_p solve $\sigma^2(\lambda_p) = \sigma_*^2 - p(1-p)$, possible for $0 \le p \le 1$, since $\sigma_*^2 \ge 1/4$ and the function σ^2 has an inverse; note also that $\lambda_0 = \lambda_1$. Define $m_p := \mu_* - \mu(\lambda_p) - p$, continuous under the assumptions on σ^2, and observe that $m_0 = m_1 + 1$. Hence the value $m = \lfloor m_0 \rfloor$ is realized in the form m_p for some $0 \le p < 1$, and then λ_p, m_p and p satisfy (4.8). In the Poisson case, for instance, this gives

$$\begin{aligned} m &= \lfloor \tilde a_1^2 - 2\tilde a_2 \rfloor; \quad p^2 \ = \ \langle \tilde a_1^2 - 2\tilde a_2 \rangle; \\ \lambda' &= \lambda + 2\tilde a_2 + \tilde a_1(1-\tilde a_1) - p(1-p), \end{aligned} \tag{4.9}$$

where $\langle x \rangle$ denotes the fractional part of x.

Suppose now that we have an approximation of a distribution P_X by some measure $\nu_r(R_\lambda; \tilde a_1, \dots, \tilde a_r)$, for $r \ge 2$, and with $\rho(\lambda) \ge 1$. We wish to show that $Q_{mp}(R_{\lambda'})$ and $\nu_r = \nu_r(R_\lambda; \tilde a_1, \dots, \tilde a_r)$ are close to order $O\{\rho(\lambda)^{-3/2}\}$, if λ', m and p are suitably chosen. Matching the first two moments, the choices of λ', m and p in (4.8) when μ_* and σ_*^2 are given by (4.5) are such as to give

$$\mu(\lambda') + m + p \ = \ \mu(\lambda) + \tilde a_1; \quad \sigma^2(\lambda') + p(1-p) \ = \ \sigma^2(\lambda) + 2\tilde a_2 + \tilde a_1(1-\tilde a_1),$$

implying that, for $b_1 := b_1(\lambda, \lambda')$ and $b_2 := b_2(\lambda, \lambda')$ given in (4.3),

$$b_1 + m + p \ = \ \tilde a_1; \qquad b_2 + (m+p)b_1 + mp + \binom{m}{2} \ = \ \tilde a_2; \tag{4.10}$$

note also that m, $|\lambda' - \lambda|$, $|\mu(\lambda') - \mu(\lambda)|$ and $|\sigma^2(\lambda') - \sigma^2(\lambda)|$ are uniformly bounded for $(\tilde a_1, \tilde a_2)$ in any compact set. Now, from the definition of $\delta_{\lambda,\lambda'}$ and from (4.7), $q_{mp}(R_{\lambda'})(\theta)$ can be written as $r_\lambda(\theta)\psi_{\lambda,\lambda'}(\theta)$, with

$$\psi_{\lambda,\lambda'}(\theta) \ = \ \{\delta_{\lambda,\lambda'}(\theta) + [1 + b_1(e^{i\theta}-1) + b_2(e^{i\theta}-1)^2]\}e^{i\theta m}(1 + p(e^{i\theta}-1)). \tag{4.11}$$

However, in view of (4.10),

$$(1 + b_1 w + b_2 w^2)(1+w)^m(1+pw) - (1 + \tilde a_1 w + \tilde a_2 w^2)$$

is a polynomial in w that begins with the w^3-term, so that

$$\hat d(\theta) \ := \ [1 + b_1(e^{i\theta}-1) + b_2(e^{i\theta}-1)^2]e^{i\theta m}(1 + p(e^{i\theta}-1)) - \tilde\psi_r(\theta) \tag{4.12}$$

satisfies

$$\hat{d}(0) = \hat{d}'(0) = 0; \quad |\hat{d}''(\theta)| \le \hat{\gamma}|\theta|, \text{ in } |\theta| \le \pi, \tag{4.13}$$

where $\hat{\gamma} = \hat{\gamma}(\tilde{a}_1, \ldots, \tilde{a}_r)$ remains bounded if $\tilde{a}_1, \ldots, \tilde{a}_r$ do. In view of (4.4) and (4.11)–(4.13), q_{mp} and $\tilde{\psi}_r r_\lambda$ are $(0, 0, \pi)$-smoothly mod r_λ polynomially close, with $M = 1$ and $t_1 = 3$, for a constant $\gamma_1 = \gamma_1(\tilde{a}_1, \ldots, \tilde{a}_r)$, whose definition depends on the family R_λ. In view of Corollary 4.1(5), this proves the following result.

Proposition 4.2 *Suppose that the family $\{R_\lambda, \lambda > 0\}$ is amenable, and that λ', m and p are chosen to satisfy (4.8) for μ_* and σ_*^2 given by (4.5). Then*

$$\|Q_{mp}(R_{\lambda'}) - \nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)\| \le \alpha' \gamma_1 \rho(\lambda)^{-3/2}. \tag{4.14}$$

Thus the signed measure $\nu_r(R_\lambda; \tilde{a}_1, \ldots, \tilde{a}_r)$ can be replaced as approximation by the probability distribution $Q_{mp}(R_{\lambda'})$ with an additional error in total variation of order at most $O(\rho(\lambda)^{-3/2})$.

Suppose that, instead of having a bound on $d_{\mu\nu} := \psi - \tilde{\psi}_r$, we are given an approximation to ψ by a Taylor expansion $\psi_r(\theta) := \sum_{l=0}^r a_l (i\theta)^l$ around $\theta = 0$, for real coefficients a_l (and with $a_0 = 1$) and some $r \in \mathbb{N}_0$. Then, equating coefficients of $i\theta$, it follows that

$$|\psi_r(\theta) - \tilde{\psi}_r(\theta)| \le U_r |\theta|^{r+1}, \quad |\theta| \le \pi, \tag{4.15}$$

for $U_r := U_r(a_1, \ldots, a_r)$, if $\tilde{a}_1, \ldots, \tilde{a}_r$ are defined implicitly by

$$a_j := \sum_{l=1}^{j} \tilde{a}_l \sum_{(s_1,\ldots,s_l)\in S_{j-l}} \prod_{t=1}^{l} \frac{1}{(s_t+1)!}, \tag{4.16}$$

where $S_m := \{(s_1, \ldots, s_l)\colon \sum_{t=1}^l s_t = m\}$. Hence we can replace any bound on the difference $\psi - \psi_r$ by a corresponding bound on $d_{\mu\nu}$ in the assumptions of the theorems, in which the original bound is increased by $U_r|\theta|^{r+1}$. This will typically not change the order of the approximation obtained.

Sometimes it is convenient, for simplicity, to use parameters in the expansions that are not those emerging naturally from the proofs. Under the conditions on the family $\{R_\lambda, \lambda > 0\}$ imposed in this section, this is easy to accommodate. For instance, suppose that, for $|\theta| \le \pi$,

$$\phi_\mu := r_\lambda A; \quad \phi_{\nu^{(1)}} := r_\lambda A'; \quad \phi_{\nu^{(2)}} := r_{\lambda'} A,$$

with $A(\theta) := 1 + \sum_{l=1}^r a_l (e^{i\theta} - 1)^l$, $A'(\theta) := 1 + \sum_{l=1}^r a'_l (e^{i\theta} - 1)^l$ and with $\lambda > \lambda'$. Then $d_{\mu\nu}^{(1)} := A - A'$ satisfies

$$|d_{\mu\nu}^{(1)}(\theta)| \le \sum_{l=1}^{r} |a_l - a'_l|\, |\theta|^l, \quad 0 < |\theta| \le \pi,$$

enabling ϕ_μ to be replaced by $\phi_{\nu^{(1)}}$ in exchange for an error that can be bounded using Corollary 4.1. Similarly, setting $d^{(2)}_{\mu\nu} := A(1 - r_{\lambda'}/r_\lambda)$, we have

$$|d^{(2)}_{\mu\nu}(\theta)| \;\le\; \tilde{f}(|\lambda - \lambda'|)|\theta| \left\{ 1 + \sum_{l=1}^{r} |a_l|\,|\theta|^l \right\}, \quad 0 < |\theta| \le \pi,$$

in view of (4.2), where $\tilde{f}\colon\ \mathbb{R}_+ \to \mathbb{R}_+$ is continuous.

5 Poisson–Charlier expansions

As observed above, the Poisson family satisfies all the requirements placed on the family $\{R_\lambda,\ \lambda > 0\}$ in the previous section, so all the results of that section can be carried across. In this case, the signed measures ν_r on $\mathbb{N}_0$ have the explicit representation

$$\nu_r\{j\} \;:=\; \nu_r(\mathrm{Po}\,(\lambda); \tilde{a}_1, \ldots, \tilde{a}_r)\{j\} \;:=\; \mathrm{Po}\,(\lambda)\{j\} \left\{ 1 + \sum_{l=1}^{r} (-1)^l \tilde{a}_l C_l(j;\lambda) \right\}, \tag{5.1}$$

where

$$C_l(j;\lambda) \;:=\; \sum_{k=0}^{l} (-1)^k \binom{l}{k} \binom{j}{k} k!\, \lambda^{-k} \tag{5.2}$$

denotes the l-th Charlier polynomial [2, (1.9), p. 171].

Note that, if $\binom{j}{k}$ is replaced by $j^k/k!$ in (5.2), one obtains the binomial expansion of $(1 - j/\lambda)^l$. As this suggests, the values of $C_l(j;\lambda)$ are in fact small for j near λ if λ is large:

$$|C_l(j;\lambda)| \;\le\; 2^{l-1}\{|1 - j/\lambda|^l + (l/\sqrt{\lambda})^l\} \tag{5.3}$$

[1, Lemma 6.1]. The equation (5.3) thus implies that, in any interval of the form $|j - \lambda| \le c\sqrt{\lambda}$, which is where the probability mass of Po (λ) is mostly to be found, the correction to the Poisson measure Po (λ) is of uniform relative order $O(\lambda^{-l/2})$. Indeed, the Chernoff inequalities for $Z \sim$ Po (λ) can be expressed in the form

$$\begin{aligned} &\max\{\mathbb{P}[Z > \lambda(1+\delta)], \mathbb{P}[Z < \lambda(1-\delta)]\} \\ &\quad \le\; \exp\{-\lambda\delta^2/2(1+\delta/3)\} \;\le\; \exp\{-\lambda\delta^2/3(\delta \vee 1)\}, \end{aligned} \tag{5.4}$$

[3, Theorem 3.2]. Since also, from (5.2),

$$|C_l(j;\lambda)| \;\le\; (1 + j/\lambda)^l \;\le\; 2^l \quad \text{if } 0 \le j \le \lambda,$$

and since

$$\binom{j}{k} k!\,\lambda^{-k}\,\frac{e^{-\lambda}\lambda^{j}}{j!} \;=\; \frac{e^{-\lambda}\lambda^{j-k}}{(j-k)!} \;\le\; \frac{e^{-\lambda}\lambda^{j-l}}{(j-l)!}$$

if $0 \le k \le l$ and $j \ge l + \lambda$, it follows that, for any $l \ge 0$, we have

$$\sum_{j=0}^{m} |C_l(j;\lambda)| \mathrm{Po}\,(\lambda)\{j\} \;\le\; 2^l \mathbb{P}[Z \le m] \;\le\; 2^l \exp\{-(\lambda - m)^2/3\lambda\}$$

for $m \le \lambda$, and, for $l \le r$ and $m \ge \lambda + r$,

$$\begin{aligned}\sum_{j \ge m} |C_l(j;\lambda)| \mathrm{Po}\,(\lambda)\{j\} \;&\le\; 2^l \mathbb{P}[Z \ge m - l] \;\le\; 2^l \mathbb{P}[Z \ge m - r] \\ &\le\; 2^l \left\{\exp\{-(m - r - \lambda)^2/3\lambda\} \vee \exp\{-(m - r - \lambda)/3\}\right\}.\end{aligned}$$

It thus follows that

$$\begin{aligned}|\nu_r|\{[0, m]\} &\le \bar{A}_r\, e^{-(\lambda - m)^2/3\lambda}, \quad 0 \le m \le \lambda; \\ |\nu_r|\{[m, \infty)\} &\le \bar{A}_r \left\{e^{-(m-r-\lambda)^2/3\lambda} \vee e^{-(m-r-\lambda)/3}\right\}, \quad m \ge \lambda + r, \end{aligned} \tag{5.5}$$

where $\bar{A}_r := 1 + \sum_{l=1}^{r} 2^l |\tilde{a}_l|$, demonstrating exponential concentration of measure for ν_r on a scale of $\sqrt{\lambda}$ around λ. Moreover, it can be deduced from (5.3) that there exists a positive constant $d = d(\tilde{a}_1, \ldots, \tilde{a}_r)$ such that $\nu_r\{j\} \ge 0$ for $|j - \lambda| \le d\lambda$, and it follows from (5.5) that $|\nu_r|\{j\colon\; |j - \lambda| > d\lambda\} = O(e^{-\alpha\lambda})$ for some $\alpha > 0$. Since also $\nu_r\{\mathbb{N}_0\} = 1$, it thus follows that, even if ν_r is formally a signed measure, it differs from a probability only on a set of measure exponentially small with λ.

Since the measures ν_r are so well concentrated, the bounds in Corollary 4.1(2–4) can be made more specific. We give as example a theorem deriving from Part 3, under the simplest conditions.

Theorem 5.1 *Suppose that X is as above, having characteristic function $\phi_X := \psi p_\lambda$, and that (2.5) holds; write $t = r + \delta$.*

If $\lambda \ge 1$, there is a constant $\alpha_{4t} = \alpha_{4t}(\tilde{a}_0, \ldots, \tilde{a}_r)$ such that

$$\|P_X - \nu_r\| \;\le\; \alpha_{4t} K_{r\delta} \lambda^{-t/2} \max\left\{1, \sqrt{|\log K_{r\delta}|}, \sqrt{\log(\lambda + 1)}\right\}; \tag{5.6}$$

if $\lambda < 1$, then there is a constant $\alpha_{5t} = \alpha_{5t}(\tilde{a}_0, \ldots, \tilde{a}_r)$ such that

$$\|P_X - \nu_r\| \;\le\; \alpha_{5t} K_{r\delta} \lambda^{-t/2} \max\{1, |\log K_{r\delta}|\}. \tag{5.7}$$

Remark Of course, for the bound in (5.7) to be of use, $K_{r\delta}$ should be small.

Proof For $\lambda \geq 1$, we use both parts of (5.5), taking

$$a := \lfloor \lambda - c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \rfloor \quad \text{and} \quad b := \lceil \lambda + r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \rceil,$$

where $\lfloor x \rfloor \leq x \leq \lceil x \rceil$ denote the integers closest to x, and with

$$c_{r\lambda} := 3\{(r+1)/2 + |\log K_{r\delta}|/\log(\lambda+1)\}.$$

If $r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} \leq \lambda$, we obtain

$$|\nu_r|\{[a,b]^c\} \leq 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}^2/3} \leq 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}/3},$$

since $c_{r\lambda} \geq 1$, and, if $r + c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} > \lambda$, we get

$$|\nu_r|\{[a,b]^c\} \leq 2\bar{A}_r \exp\{-c_{r\lambda}\sqrt{\lambda \log(\lambda+1)}/3\} \leq 2\bar{A}_r(\lambda+1)^{-c_{r\lambda}/3},$$

since $\lambda \geq \log(\lambda+1)$ in $\lambda \geq 0$. Hence, in either case, from the definition of $c_{r\lambda}$, we have

$$|\nu_r|\{[a,b]^c\} \leq 2\bar{A}_r K_{r\delta}(\lambda+1)^{-(r+1)/2}. \tag{5.8}$$

Hence, from Corollary 4.1(3), with $\varepsilon = \eta = 0$, $M = 1$, $\gamma_1 = K_{r\delta}$ and $t_1 = t$, it follows that

$$\|P_X - \nu_r\| \leq \left\{2c_{r\lambda}\sqrt{\lambda \log(\lambda+1)} + r + 2\right\}\alpha'_{1t}K_{r\delta}\lambda^{-(t+1)/2} \\ + \alpha'_{2t}K_{r\delta}\lambda^{-t/2} + 4\bar{A}_r K_{r\delta}\lambda^{-(r+1)/2},$$

with

$$\alpha'_{1t} := \alpha_{1t}(\pi^2/2)^{(t+1)/2}; \quad \alpha'_{2t} := \alpha_{2t}(\pi^2/2)^{t/2}, \tag{5.9}$$

so that

$$\|P_X - \nu_r\| \leq \beta_{3t}K_{r\delta}\lambda^{-t/2}\sqrt{\log(\lambda+1)}\max\left\{1, \frac{|\log K_{r\delta}|}{\log(\lambda+1)}\right\},$$

with $\beta_{3t} := \alpha'_{1t}\{4r+11\} + \alpha'_{2t} + 4\bar{A}_r$.

For $\lambda < 1$, we take $b := \lceil 2 + r + 3|\log K_{r\delta}|\rceil$ in (5.5), giving

$$|\nu_r|\{[b,\infty)\} \leq \bar{A}_r K_{r\delta},$$

and then, from Corollary 4.1(3) as above, it follows that

$$\|P_X - \nu_r\| \leq (r + 3 + 3|\log K_{r\delta}|)\alpha'_{1t}K_{r\delta} + \alpha'_{2t}K_{r\delta} + 2\bar{A}_r K_{r\delta},$$

so that

$$\|P_X - \nu_r\| \;\le\; \beta'_{3t} K_{r\delta} \max\{1, |\log(K_{r\delta}|\},$$

with $\beta'_{3t} := \alpha'_{1t}\{r+6\} + \alpha'_{2t} + 2\bar{A}_r$. □

6 Compound Poisson approximation

The theory of Sect. 3 can also be applied when the distributions R_λ come from a compound Poisson family. For $\lambda > 0$ and for μ a probability distribution on $\mathbb{Z}$, let CP (λ, μ) denote the distribution of the sum $Y := \sum_{j\in\mathbb{Z}\setminus\{0\}} jZ_j$, where $Z_j,\ j \ne 0$, are independent, and $Z_j \sim$ Po $(\lambda\mu_j)$. Then, if $\mu_1 > 0$, the characteristic function of Y is of the form $R_\lambda := \zeta_\lambda p_{\lambda_1}$, where ζ_λ is the characteristic function of $\sum_{j\in\mathbb{Z}\setminus\{0,1\}} jZ_j$ and $\lambda_1 = \lambda\mu_1$. Thus, for the purposes of applying Corollary 4.1, ρ can be taken to be $2\pi^{-2}\mu_1\lambda$.

These considerations apply as long as $\mu_1 > 0$, and could also be invoked if $\mu_{-1} > 0$. If $\mu_1 = \mu_{-1} = 0$, there is then no factor of the form p_λ to guarantee that, for some $\rho > 0$, the characteristic function ϕ_Y of Y is (ρ, π)-locally normal. Indeed, if $Y = 2Z$ where $Z \sim$ Po (λ), and if $W \sim$ Be $(1/2)$ is independent of Y, it is not true that the distribution of $Y + W$ is close to that of Y in total variation, even though $|\phi_{Y+W}(\theta) - \phi_Y(\theta)| \le K_0|\theta|\,|\phi_Y(\theta)|$; this is to be compared to the example in Sect. 2.

7 Applications

7.1 Sums of independent random variables

Let $X_1, \ldots, X_n$ be independent integer valued random variables, and let S_n denote their sum. In contexts in which a central limit approximation to the distribution of S_n would be appropriate, the classical Edgeworth expansion (see, e.g., [7, Chapter 5]) is unwieldy, because S_n is confined to the integers. As an alternative, Barbour and Čekanavičius [1, Theorem 5.1] give a Poisson–Charlier expansion, for S_n 'centred' so that its mean and variance are almost equal. If the X_i have variances that are uniformly bounded below and have bounded $(r+1+\delta)$-th moments, and if the distribution of each X_i has non-trivial overlap with that of $X_i + 1$, their error bound with respect to the total variation norm is of order $O(n^{-(r-1+\delta)/2})$. Here, under similar conditions, we use Corollary 4.1 to prove an error bound for their expansion which is of the same order, but is established only with respect to the less stringent Kolmogorov distance. A total variation bound for the error, of the slightly larger order $O(n^{-(r-1+\delta)/2}\sqrt{\log n})$, could be deduced from Corollary 4.1(3), by taking $a = \lfloor \lambda - k\sqrt{\lambda\log\lambda}\rfloor$ and $b = \lceil \lambda + k\sqrt{\lambda\log\lambda}\rceil$, for suitable choice of $k = k_r$, where $\lambda = \mathbb{E}S_n$ (and $\mathbb{E}S_n \approx \mathrm{Var}\, S_n$, because of centring).

Assume that each of the X_j has finite $(r+1+\delta)$-th moment, with $r \geq 1$, and define

$$A^{(r)}(w) \;:=\; 1+\sum_{l\geq 2}\tilde{a}_l^{(r)}w^l \;=\; \exp\left\{\sum_{l=2}^{r+1}\frac{\kappa_l w^l}{l!}\right\}, \tag{7.1}$$

where $\kappa_l := \kappa_l(S_n)$ and $\kappa_l(X)$ denotes the l-th factorial cumulant of the random variable X. Then the approximation that we establish is to the Poisson–Charlier signed measure ν_r with

$$\nu_r\{j\} \;:=\; \mathrm{Po}\,(\lambda)\{j\}\left\{1+\sum_{l=2}^{L_r}(-1)^l\tilde{a}_l^{(r)}C_l(j;\lambda)\right\}, \tag{7.2}$$

where $L_r := \max\{1, 3(r-1)\}$, and where $\lambda := \mathbb{E}S_n$; ν_r has characteristic function

$$\phi_{\nu_r} \;:=\; p_\lambda(\theta)\,\widetilde{A}^{(r)}(\theta), \tag{7.3}$$

where

$$\widetilde{A}^{(r)}(\theta) \;:=\; 1+\sum_{l=2}^{L_r}\tilde{a}_l^{(r)}(e^{i\theta}-1)^l. \tag{7.4}$$

We need two further quantities involving the X_j:

$$K^{(n)} \;:=\; \left|\sum_{j=1}^{n}\kappa_2(X_j)\right| \;=\; |\mathrm{Var}\,S_n-\mathbb{E}S_n|, \tag{7.5}$$

kept small by judicious centring, and

$$p_j \;:=\; 1-\tfrac{1}{2}\|\mathcal{L}(X_j)-\mathcal{L}(X_j+1)\|. \tag{7.6}$$

Theorem 7.1 *Suppose that there are constants K_l, $1 \leq l \leq r+1$, such that, for each j,*

$$|\kappa_l(X_j)| \;\leq\; K_l, \quad 2\leq l\leq r+1; \qquad \mathbb{E}|X_j|^{r+1+\delta} \;\leq\; K_1^{r+1+\delta}.$$

Suppose also that $p_j \geq p_0 > 0$ for all j, and that $\lambda \geq n\lambda_0$. Then

$$d_{\mathrm{K}}(\mathcal{L}(S_n),\nu_r) \;\leq\; G(K_1,\ldots,K_{r+1},K^{(n)},p_0^{-1},\lambda_0^{-1})n^{-(r-1+\delta)/2},$$

for a function G that is bounded on compact sets.

Remark For asymptotics in n, with triangular arrays of variables, the error is of order $O(n^{-(r-1+\delta)/2})$ when λ_0 and p_0 are bounded away from zero, and $K_1, \ldots, K_{r+1}$ and $K^{(n)}$ remain bounded. The requirements on λ_0 and p_0 can often be achieved by grouping the random variables appropriately, though attention then has to be paid to the consequent changes in the K_l. The condition (7.5) can always be satisfied with $K^{(n)} \leq 1$, by replacing the X_j by translates, where necessary. For more discussion, we refer to [1]. The above conditions are designed to cover sums of independent random variables, each of which has non-trivial variance, has uniformly bounded $(r+1+\delta)$-th moment, and whose distribution overlaps with its unit translate.

Proof We check the conditions of Corollary 4.1(2). First, in view of (7.6), we can write

$$\mathbb{E}(e^{i\theta X_j}) \;=\; \tfrac{1}{2} p_j(e^{i\theta} + 1)\phi_{1j}(\theta) + (1 - p_j)\phi_{2j}(\theta),$$

where both ϕ_{1j} and ϕ_{2j} are characteristic functions. Hence we have

$$\left|\mathbb{E}(e^{i\theta X_j})\right| \;\leq\; 1 - p_j + p_j \cos(\theta/2) \;\leq\; 1 - p_j\theta^2/4\pi, \quad 0 \leq |\theta| \leq \pi.$$

Hence $\phi_\mu(\theta) := \mathbb{E}(e^{i\theta S_n})$ satisfies

$$|\phi_\mu(\theta)| \;\leq\; \exp\{-np_0\theta^2/4\pi\}, \quad 0 \leq |\theta| \leq \pi. \tag{7.7}$$

On the other hand, from the additivity of the factorial cumulants, we have

$$|\kappa_l(S_n)| \;\leq\; nK_l, \quad 3 \leq l \leq r+1,$$

with $|\kappa_2(S_n)| \leq K^{(n)}$ from (7.5). From (7.1), we thus deduce the bound $|\tilde{a}_l^{(r)}| \leq c_l n^{\lfloor l/3 \rfloor}$, for $c_l = c_l(K^{(n)}, K_3, \ldots, K_{r+1})$, $l \geq 1$. Hence

$$|\phi_{\nu_r}(\theta)| \;\leq\; \exp\{-2n\lambda_0\theta^2/\pi^2\}c' n^{\lfloor L_r/3 \rfloor} \;\leq\; \exp\{-n\lambda_0\theta^2/\pi^2\}c'', \tag{7.8}$$

for $c'' = c''(K^{(n)}, K_3, \ldots, K_{r+1})$. Combining (7.7) and (7.8), we can thus take $\eta := Ce^{-n\rho'\theta_0^2}$ in (3.2), for

$$\rho' = \min\{\lambda_0/\pi^2, p_0/4\pi\}$$

and a suitable $C = C(K^{(n)}, K_3, \ldots, K_{r+1})$. The choice of θ_0 we postpone for now.

For $|\theta| \leq \theta_0$, we take $\chi(\theta) := p_\lambda(\theta)$, and check the approximation of

$$\psi_\mu(\theta) \;:=\; \phi_\mu(\theta)\exp\{-\lambda(e^{i\theta} - 1)\} \;=\; \mathbb{E}\left\{(1+w)^{S_n}\right\} e^{-w\mathbb{E}S_n}$$

by $\widetilde{A}^{(r)}(\theta)$ as a polynomial in $w := e^{i\theta} - 1$. We begin with the inequality

$$\begin{aligned} \left|(1+w)^s - \sum_{l=0}^{r+1} \frac{w^l}{l!}\, s_{(l)}\right| &\le \frac{|s_{(r+2)}|}{(r+2)!}\,|w|^{r+2} \wedge 2\frac{|s_{(r+1)}|}{(r+1)!}\,|w|^{r+1} \\ &\le \frac{|s_{(r+1)}|}{(r+2)!}\,|w|^{r+1+\delta}\{|s|+r+1\}^{\delta}\{2(r+2)\}^{1-\delta}, \end{aligned}$$

derived using Taylor's expansion, true for any $s \in \mathbb{Z}$ and $0 < \delta \le 1$, where $s_{(l)} := s(s-1)\dots(s-l+1)$. Hence, for each j, we have

$$\left|\mathbb{E}\left\{(1+w)^{X_j}\right\} - \sum_{l=0}^{r+1} \frac{\mathbb{E}\{(X_j)_{(l)}\}}{l!}\, w^l\right| \le c_{r,\delta}|\theta|^{r+1+\delta}(K_1 + K_1^{r+1+\delta}), \quad (7.9)$$

for a universal constant $c_{r,\delta}$. Then, writing

$$Q_{r+1}^{(s)}(w; X) := \exp\left\{\sum_{l=s}^{r+1} \kappa_l(X) w^l/l!\right\},$$

and using the differentiation formula in [7, p. 170], we have

$$\begin{aligned} &\left|Q_{r+1}^{(1)}(w; X_j) - \sum_{l=0}^{r+1} \frac{\mathbb{E}\{(X_j)_{(l)}\}}{l!}\, w^l\right| \\ &\le \frac{|\theta|^{r+2}}{(r+2)!} \sup_{|\theta'|\le\theta_0} \left|\frac{d^{r+2}}{dz^{r+2}} Q_{r+1}^{(1)}(z; X_j)\right|_{z=e^{i\theta'}-1} \\ &\le |\theta|^{r+2} c(K_1, \dots, K_{r+1}), \end{aligned} \quad (7.10)$$

for a suitable function c and for all $|\theta| \le \pi$. Combining these estimates, we deduce that, for $w = e^{i\theta} - 1$ and for all $|\theta| \le \pi$,

$$\left|\mathbb{E}\left\{(1+w)^{X_j}\right\} e^{-\mathbb{E}X_j w} - Q_{r+1}^{(2)}(w; X_j)\right| \le k_1|\theta|^{r+1+\delta}, \quad (7.11)$$

where $k_1 = k_1(K_1, \dots, K_{r+1})$.

Now a standard inequality shows that, for $u_j := \prod_{l=1}^{j} x_l \prod_{l=j+1}^{n} y_l$, for complex x_l, y_l with $y_l \ne 0$ and $|x_l/y_l - 1| \le \varepsilon_l$, then

$$|u_n - u_0| \le |u_0| \left\{\prod_{s=1}^{n-1}(1+\varepsilon_s)\right\} \sum_{l=1}^{n} \varepsilon_l. \quad (7.12)$$

Taking $x_j := \mathbb{E}\{(1+w)^{X_j}\} + e^{-\mathbb{E}X_j w}$ and $y_j := Q^{(2)}_{r+1}(w; X_j)$, (7.11) shows that we can take $\varepsilon_l := \varepsilon := k_1|\theta|^{r+1+\delta}e^E$ for each l, with

$$E := \exp\left\{\sum_{l=2}^{r+1} K_l/l!\right\},$$

provided that $|\theta| \le \theta_0 \le 1$. For $r \ge 2$, choosing $\theta_0 := n^{-1/3}$ then ensures that $(1+\varepsilon)^n$ is suitably bounded, and (7.12) yields

$$\left|\mathbb{E}\left\{(1+w)^{S_n}\right\} e^{-w\mathbb{E}S_n} - Q^{(2)}_{r+1}(w; S_n)\right| \le k_2 n|\theta|^{r+1+\delta}, \tag{7.13}$$

for $k_2 = k_2(K^{(n)}, K_1, \ldots, K_{r+1})$, since

$$|u_0| := |Q^{(2)}_{r+1}(w; S_n)| \le \exp\{|\kappa_2(S_n)|\theta_0^2/2\} \exp\left\{\sum_{l=3}^{r+1} nK_l\theta_0^l/l!\right\}$$

is bounded for $\theta_0 = n^{-1/3}$, in view of (7.5). For $r = 1$, $|u_0|$ is uniformly bounded if $\theta_0 \le 1$, and the choice $\theta_0 = n^{-1/(2+\delta)}$ ensures that $(1+\varepsilon)^n$ remains bounded.

The remaining step is to note that, for $w := e^{i\theta} - 1$, $\widetilde{A}^{(r)}(\theta)$ contains all terms up to the power w^{L_r} in the power series expansion of $Q^{(2)}_{r+1}(w; S_n)$, giving

$$\left|Q^{(2)}_{r+1}(w; S_n) - \widetilde{A}^{(r)}(\theta)\right| \le \frac{|\theta|^{L_r+1}}{(L_r+1)!} \sup_{|\theta'|\le|\theta|} \left|\frac{d^{L_r+1}}{dz^{L_r+1}} Q^{(2)}_{r+1}(z; S_n)\right|_{z=e^{i\theta'}-1}. \tag{7.14}$$

Now $|\kappa_2(S_n)|$ is bounded by $K^{(n)}$, and, for $l \ge 3$, each $\kappa_l(S_n)$, for which we have only the weak bound nK_l, occurs associated with the power w^l in the exponent of $Q^{(2)}_{r+1}(w; S_n)$. Writing

$$\frac{d^s}{dz^s} Q^{(2)}_{r+1}(z; S_n) = P_s(n, z) Q^{(2)}_{r+1}(z; S_n),$$

the monomials that make up $P_s(n, z)$ thus have coefficients of magnitude n^l associated with powers z^m with $m \ge (2l - (s-l))_+ = (3l - s)_+$, so that they are themselves of magnitude at most $O(n^{l-(3l-s)_+/3}) = O(n^{s/3})$ in $|\theta'| \le n^{-1/3}$. Taking $s = L_r + 1$ and $r \ge 2$, $m = 0$ requires that $l \le r - 1$, and $l \ge r$ entails $m \ge 2$, so that, for $r \ge 2$ and $|\theta| \le \theta_0$,

$$\sup_{|\theta'|\le|\theta|} \left|\frac{d^{L_r+1}}{dz^{L_r+1}} Q^{(2)}_{r+1}(z; S_n)\right|_{z=e^{i\theta'}-1} \le k_3 n^{r-1}(1 + n|\theta|^2),$$

with $k_3 = k_3(K^{(n)}, K_1, \ldots, K_{r+1})$. If $|\theta| \ge n^{-1/2}$, it follows that the bound in (7.14) is at most $2k_3\{(L_r+1)!\}^{-1} n^r |\theta|^{3r}$; if $|\theta| \le n^{-1/2}$, the bound is at most

$2k_3\{(L_r+1)!\}^{-1}n|\theta|^{r+2}$. Combining this with (7.13), we have established that for $|\theta| \le n^{-1/3}$ and $r \ge 2$, we have

$$|\phi_\mu(\theta)\exp\{-\lambda(e^{i\theta}-1)\} - \widetilde{A}^{(r)}(\theta)| \;\le\; k_4 n|\theta|^{r+1+\delta}(1+(n|\theta|^2)^{r-1}), \tag{7.15}$$

where $k_4 = k_4(K^{(n)}, K_1, \ldots, K_{r+1})$. This shows that ϕ_μ and ϕ_{ν_r} are $(0,\eta,\theta_0)$-mod p_λ polynomially close, with

$$M=2, \quad \gamma_1 = nk_4, \quad t_1 = r+1+\delta, \quad \gamma_2 = n^r k_4, \quad t_2 = 3r-1+\delta,$$

and with $\theta_0 = n^{-1/3}$ and $\eta = Ce^{-n^{1/3}\rho'}$, this last from the bounds (7.7) and (7.8). Applying Corollary 4.1(2), taking $a=0$ and $b=2\lambda$, and using the tail properties of the Poisson–Charlier measures (5.5), the theorem follows for $r \ge 2$.

For $r=1$, the bound in (7.14) is easily of order $|\theta|^2$, giving a bound in (7.15) of $k_4'(n|\theta|^{2+\delta} + |\theta|^2)$. This leads to the choices

$$M=2, \quad \gamma_1 = nk_4', \quad t_1 = 2+\delta, \quad \gamma_2 = k_4', \quad t_2 = 2d,$$

together with $\theta_0 = n^{-1/(2+\delta)}$ and $\eta = Ce^{-n^{\delta/(2+\delta)}\rho'}$, and the remainder of the proof is as before. □

7.2 Analytic combinatorial schemes

An extremely interesting range of applications is to be found in the paper of Hwang [5]. His conditions are motivated by examples from combinatorics, in which generating functions are natural tools. He works in an asymptotic setting, assuming that X_n is a random variable whose probability generating function G_n is of the form

$$G_n(z) \;=\; z^h(g(z)+\varepsilon_n(z))e^{\lambda(z-1)},$$

where h is a non-negative integer, and both g and ε_n are analytic in a closed disc of radius $\eta > 1$. As $n \to \infty$, he assumes that $\lambda \to \infty$ and that $\sup_{z:|z|\le\eta} |\varepsilon_n(z)| \le K\lambda^{-1}$, uniformly in n. He then proves a number of results describing the accuracy of the approximation of $P_{X_n - h}$ by Po $(\lambda + g'(1))$.

Under his conditions, it is immediate that we can write

$$g(z) \;=\; \sum_{j\ge0} g_j(z-1)^j \quad \text{and} \quad \varepsilon_n(z) \;=\; \sum_{j\ge0} \varepsilon_{nj}(z-1)^j \tag{7.16}$$

for $|z-1| < \eta$, with

$$|g_j| \;\le\; k_g(\eta-1)^{-j} \quad \text{and} \quad |\varepsilon_{nj}| \;\le\; \lambda^{-1}k_\varepsilon(\eta-1)^{-j} \tag{7.17}$$

for all $j \geq 0$. Hence $X := X_n - h$ has characteristic function of the form $\psi^{(n)} p_\lambda$, where

$$\psi^{(n)}(\theta) = g(e^{i\theta}) + \varepsilon_n(e^{i\theta}),$$

and thus, for any $r \in \mathbb{N}_0$,

$$|\psi^{(n)}(\theta) - \tilde{\psi}_r^{(n)}(\theta)| \leq K_{r1}|\theta|^{r+1}, \quad |\theta| \leq (\eta - 1)/2, \tag{7.18}$$

with $\tilde{\psi}_r^{(n)}$ defined as in (2.4), taking $\tilde{a}_j^{(n)} = g_j + \varepsilon_{nj}$; note that the constant K_{r1} can indeed be taken to be uniform for all n. Since also g and ε_n are both uniformly bounded on the unit circle, and since $\tilde{\psi}_r^{(n)}$ is bounded (uniformly in n) for $|\theta| \leq \pi$, it is clear that (7.18) can be extended to all $|\theta| \leq \pi$, albeit with a different uniform constant K'_{r1}, so that (2.5) holds with $\delta = 1$ for any $r \in \mathbb{N}_0$. Thus Parts 1–3 of Corollary 4.1 (with $R_\lambda = \mathrm{Po}(\lambda)$ and $\rho(\lambda) = 2\lambda/\pi^2$) can be applied with any choice of r, giving progressively more accurate approximations to P_{X_n-h}, as far as the λ-order is concerned, in terms of progressively more complicated perturbations of the Poisson distribution. These theorems are thus applicable to all the examples that Hwang considers, including the numbers of components (counted in various ways) in a wide class of logarithmic assemblies, multisets and selections.

For instance, using translated Poisson approximation as in Sect. 4.2 by way of Proposition 4.2 gives an approximation to P_{X_n-h} by the mixture $Q_{mp}(\mathrm{Po}(\lambda'))$, where, from (4.9),

$$m := \lfloor m_n - v_n \rfloor; \quad p^2 := \langle m_n - v_n \rangle; \quad \lambda' := \lambda + v_n - p(1-p),$$

where $m_n := g'_n(1)$, $v_n := g''_n(1) + g'_n(1) - \{g'_n(1)\}^2$ and $g_n := g + \varepsilon_n$. Hwang's approximation by $\mathrm{Po}(\lambda + g'(1))$ has asymptotically the same mean as ours (and as that of $X_n - h$), but a variance asymptotically differing by $\kappa := g''(1) - \{g'(1)\}^2$. As a consequence, Hwang's approximation has an error of larger asymptotic order, in which the quantity κ appears; for instance, for Kolmogorov distance, his Theorem 1 gives an error of order $O(\lambda^{-1})$, whereas that obtained using Corollary 4.1(2) together with Proposition 4.2 is of order $O(\lambda^{-3/2})$.

Although our Poisson expansion theorems are automatically applicable under Hwang's conditions, they also apply to examples that do not satisfy his conditions: the simple example at the end of Sect. 2 is one such. Conversely, Hwang's Theorem 2, which establishes Poisson approximation in the lower tail with good *relative* accuracy, cannot be proved using only our conditions; the conclusion would not be true, for instance, in the example just mentioned.

Note also that Hwang examines problems from combinatorial settings in which approximation is not by Poisson distributions: he has examples concerning the (amenable) Bessel family of distributions,

$$B(\lambda)\{j\} := L(\lambda)^{-1} \frac{\lambda^j}{j!(j-1)!}, \quad j \in \mathbb{N},$$

for the appropriate choice of normalizing constant $L(\lambda)$. Thus we could apply Corollary 4.1 to obtain asymptotically more accurate expansions, and, in conjunction with Proposition 4.2, obtain slightly sharper approximations than his within the translated Bessel family.

7.3 Prime divisors

The numbers of prime divisors of a positive integer n, counted either with ($\Omega(n)$) or without ($\omega(n)$) multiplicity, can also be treated by these methods, since excellent information is available about their generating functions. For our purposes, we use only the shortest expansion, taken from [11, Theorems II.6.1 and 6.2]. One finds that, for N_n uniformly distributed on $\{1, 2, \ldots, n\}$, the characteristic functions of $\Omega(n)$ and $\omega(n)$ are given by

$$\mathbb{E}\{e^{i\theta\omega(N_n)}\} = p_{\log\log n}(\theta)\left\{\Phi_1(e^{i\theta}-1)+\varepsilon_1(\theta)\right\};$$
$$\mathbb{E}\{e^{i\theta\Omega(N_n)}\} = p_{\log\log n}(\theta)\left\{\Phi_2(e^{i\theta}-1)+\varepsilon_2(\theta)\right\},$$

where $|\varepsilon_s(\theta)| \le C_s/\log n,\ s = 1, 2$, for some constants C_1 and C_2, and

$$\Phi_1(w) := \frac{1}{\Gamma(1+w)}\prod_q\left(1+\frac{w}{q}\right)\left(1-\frac{1}{q}\right)^w;$$
$$\Phi_2(w) := \frac{1}{\Gamma(1+w)}\prod_q\left(1-\frac{w}{q-1}\right)^{-1}\left(1-\frac{1}{q}\right)^w,$$

q running here over prime numbers. These expansions were established and used by Rényi and Turán [9] in their proof of the Erdős–Kac Theorem, but they are also sketched by Selberg [10].

Kowalski and Nikeghbali [6] have emphasized the structural interpretation of these functions, which we now recall. Write

$$\Phi_{1,1}(\theta) = \frac{1}{\Gamma(e^{i\theta})}, \quad \Phi_{1,2}(\theta) = \prod_q\left(1+\frac{e^{i\theta}-1}{q}\right)\left(1-\frac{1}{q}\right)^{e^{i\theta}-1},$$

so that $\Phi_1(e^{i\theta}-1) = \Phi_{1,1}(\theta)\Phi_{1,2}(\theta)$.

Let X_n be the random variable giving the number of disjoint cycles appearing in the decomposition of a random uniformly distributed permutation of size n. In addition, let Y_n be a random variable of the form

$$Y_n = \sum_{q\le n} B_q$$

where the B_q are independent Bernoulli random variables indexed by primes, with $\mathbb{P}[B_q = 1] = 1/q$; Y_n represents a naive model of the number of prime divisors $\le n$ of a large integer.

Then we have

$$\mathbb{E}\{e^{i\theta X_n}\} \sim p_{\log n}(\theta)\Phi_{1,1}(\theta),$$

and

$$\mathbb{E}\{e^{i\theta Y_n}\} \sim p_{\log\log n}(\theta)\Phi_{1,2}(\theta).$$

This suggests an interpretation of the Rényi–Turán formula as a probabilistic decomposition of $\omega(N_n)$ in terms of random permutations of size $\log n$ and the naive divisibility model for integers, with an intricate dependency structure. We note that in the setting of polynomials over finite fields, this interpretation was shown by Kowalski and Nikeghbali [6] to have a precise meaning and to be very useful.

We come back to the application of our results to $\omega(N_n)$ and $\Omega(N_n)$. Let $\tilde{a}_{ls}$, $s = 1, 2$, denote the Taylor coefficients of the functions $\Phi_s(w)$ as power series in w (around $w = 0$, which corresponds to $\theta = 0$). By analyticity near 0, it follows that, for any r, we have

$$\left|\Phi_s(w) - 1 - \sum_{l=1}^{r} \tilde{a}_{ls} w^l\right| \le C_{rs}|w|^{r+1}; \left|\Phi_s''(w) - \sum_{l=2}^{r} \tilde{a}_{ls} l(l-1)w^{l-2}\right| \le C'_{rs}|w|^{r-1},$$

for suitable constants C_{rs}, C'_{rs} and for $|w| \le 2$. In order to approximate the distributions $P_{\omega(N_n)}$ and $P_{\Omega(N_n)}$, we define the measures $\nu_r^{(s)}$ by

$$\nu_r^{(s)}\{j\} := \mathrm{Po}\,(\log\log n)\{j\}\left(1 + \sum_{l=1}^{r}(-1)^l \tilde{a}_{ls} C_l(j; \log\log n)\right),$$

and invoke Corollary 4.1 with $M = 1$, $\theta_0 = \pi$ and $\varepsilon = C_s/\log n$, together with (3.30); this leads to the following conclusion, which refines the Erdős–Kac theorem.

Theorem 7.2 *For the measures $\nu_r^{(s)}$ defined above, we have*

$$\begin{aligned}
d_{\mathrm{loc}}(P_{\omega(N_n)}, \nu_r^{(1)}) &\le \alpha'_{1,r+1} C_{r1}(\log\log n)^{-1-r/2} + \tilde{\alpha}_1 C_1/\log n;\\
\|P_{\omega(N_n)} - \nu_r^{(1)}\| &\le 2\alpha'(r+1, \pi^2/2)C'_{r1}\left(1 + \frac{2}{r}\right)(\log\log n)^{-(r+1)/2}\\
&\quad + \widetilde{C}_1 \log\log n/\log n;\\
d_{\mathrm{loc}}(P_{\Omega(N_n)}, \nu_r^{(2)}) &\le \alpha'_{1,r+1} C_{r2}(\log\log n)^{-1-r/2} + \tilde{\alpha}_1 C_2/\log n;\\
\|P_{\Omega(N_n)} - \nu_r^{(2)}\| &\le 2\alpha'(r+1, \pi^2/2)C'_{r2}\left(1 + \frac{2}{r}\right)(\log\log n)^{-(r+1)/2}\\
&\quad + \widetilde{C}_2 \log\log n/\log n,
\end{aligned}$$

for suitable constants $\widetilde{C}_1$ and $\widetilde{C}_2$, and with α'_{1l} as defined in (5.9).

Remark As far as we know, total variation approximation was first considered in this context by Harper [4], who proved a bound with error of size $1/(\log\log n)$ (for a truncated version of $\omega(n)$, counting only prime divisors of size up to $n^{1/(3(\log\log n)^2)}$), and deduced explicit bounds in Kolmogorov distance.

To indicate what this means in concrete terms for number theory readers, consider the case of $\omega(n)$ for $r = 1$. Taylor expansion gives

$$\Phi_1(w) \;=\; 1 + B_1 w + O(w^2)$$

as $w \to 0$, where $B_1 \approx 0.26149721$ is the Mertens constant, i.e., the real number such that

$$\sum_{\substack{q \le x \\ q \text{ prime}}} \frac{1}{q} = \log\log x + B_1 + o(1),$$

as $x \to +\infty$. An application of Theorem 7.2 with $r = 1$ gives

$$\begin{aligned}\left|\frac{1}{n}|\{k \le n \mid \omega(n) \in A\}| - \nu_1^{(1)}\{A\}\right| &\le \tfrac{1}{2}\|P_{\omega(N_n)} - \nu_1^{(1)}\| \\ &= O\Big(\frac{1}{\log\log n}\Big),\end{aligned}$$

for any set A of positive integers, where

$$\nu_1^{(1)}\{j\} \;=\; \mathrm{Po}\,(\log\log n)\{j\}\Big(1 - B_1\left\{1 - \frac{j}{\log\log n}\right\}\Big).$$

Higher expansions could be computed in much the same way.

Alternatively, a more accurate approximation is available from Theorem 7.2 with $r = 2$, while staying within the realm of (translated) Poisson distributions, by invoking Proposition 4.2. For this, we compute the expansion of Φ_1 to order 2, obtaining (after some calculations) that

$$\Phi_1(w) \;=\; 1 + \tilde{a}_1 w + \tilde{a}_2 w^2 + O(w^3), \qquad \text{as } w \to 0,$$

where

$$\tilde{a}_1 \;:=\; B_1; \quad \tilde{a}_2 \;:=\; \frac{B_1^2}{2} - \frac{\pi^2}{12} - \frac{1}{2}\sum_{q \text{ prime}} \frac{1}{q^2}$$

(use $1/\Gamma(1 + w) = 1 + \gamma w + (\gamma^2/2 - \pi^2/12)w^2 + O(w^3)$, as well as the Mertens identity

$$\gamma + \sum_{q \text{ prime}} \left(\frac{1}{q} + \log\left(1 - \frac{1}{q}\right)\right) = B_1,$$

and expand every term in the Euler product). This corresponds to (2.5), since $w = e^{i\theta} - 1$.

We can then apply Theorem 7.2 and Proposition 4.2 to yield the translated Poisson approximation $Q_{mp}(\mathrm{Po}\,(\lambda'))$, with λ', m and p found from (4.9). With

$$x \;:=\; \tilde{a}_1^2 - 2\tilde{a}_2 \;=\; \frac{\pi^2}{6} + \sum_{q \text{ prime}} \frac{1}{q^2} \;\approx\; 2.0971815,$$

this gives

$$p = \sqrt{\langle x \rangle} \;\approx\; 0.31173945; \quad m \;=\; 2;$$
$$\lambda' = \log\log n + B_1 - x - p(1-p) \;\approx\; \log\log n - 2.0502422.$$

Thus, for any positive integer n and any set A of positive integers, we have

$$\left| \frac{1}{n} |\{k \le n \;\mid\; \omega(n) \in A\}| - \{p\mathrm{Po}\,(\lambda')\{A-3\} + (1-p)\mathrm{Po}\,(\lambda')\{A-2\}\} \right|$$
$$= O\Big(\frac{1}{(\log\log n)^{3/2}}\Big).$$

Similar results hold for $\Omega(n)$, where one obtains the following approximate values for the quantities p, m, λ':

$$p \;\approx\; 0.5195; \quad m \;=\; 0; \quad \lambda' \;\approx\; \log\log n + 0.5152.$$

Appendix

To prove Lemma 3.4, assume without loss of generality that $a \ge 0$. If $b \ge 0$, take $w(x) = a + bx - \frac{1}{2}cx^2$ for $0 \le x \le x_1 := b/c$, when w reaches its maximum of $h := a + b^2/2c$, and continue with the same definition until $x = x_1 + x_2$, where $x_2 := \sqrt{h/c}$, at which point $w(x_1 + x_2) = h/2$. Changing the second derivative from $-c$ to c gives $w(x) = \frac{1}{2}c(x_1 + 2x_2 - x)^2$, to be used for $x_1 + x_2 \le x \le x_1 + 2x_2$, and then take $w(x) = 0$ for $x > x_1 + 2x_2$. This definition of w satisfies all the claimed requirements.

For $b < 0$, take $w(x) = a + bx + \frac{1}{2}cx^2$ until $x_1 := |b|/c$, when $w'(x_1) = 0$ and $w(x_1) = a - b^2/c$. Thereafter, continue essentially as before, with $h := |a - b^2/2c|$ and $x_2 := \sqrt{h/c}$, taking second derivative $-\mathrm{sgn}(w(x_1))c$ in $(x_1, x_1 + x_2)$ and $\mathrm{sgn}(w(x_1))c$ in $(x_1 + x_2, x_1 + 2x_2)$. □

References

1. Barbour, A.D., Čekanavičius, V.: Total variation asymptotics for sums of independent integer random variables. Ann. Probab. **30**, 509–545 (2002)
2. Chihara, T.S.: An Introduction to Orthogonal Polynomials. Gordon and Breach, New York (1978)

3. Chung, F., Lu, L.: Concentration inequalities and martingale inequalities: a survey. Int. Math. **3**, 79–127 (2006)
4. Harper, A.J.: Two new proofs of the Erdős–Kac Theorem, with bound on the rate of convergence, by Stein's method for distributional approximations. Math. Proc. Camb. Philos. Soc. **147**, 95–114 (2009)
5. Hwang, H.-K.: Asymptotics of Poisson approximation to random discrete distributions: an analytic approach. Adv. Appl. Probab. **31**, 448–491 (1999)
6. Kowalski, E., Nikeghbali, A.: Mod-Poisson convergence in probability and number theory, Int. Math. Res. Notices. doi:10.1093/imrn/rnq019; see also arXiv:0905.0318 (2010)
7. Petrov, V.V.: Limit Theorems of Probability Theory. Oxford University Press, Oxford (1975)
8. Presman, E.L.: Approximation of binomial distributions by infinitely divisible ones. Theor. Probab. Appl. **28**, 393–403 (1983)
9. Rényi, A., Turán, P.: On a theorem of Erdős–Kac. Acta Arith. **4**, 71–84 (1958)
10. Selberg, A.: Note on the paper by L.G. Sathe. J. Indian Math. Soc. **18**, 83–87 (1954)
11. Tenenbaum, G.: Introduction à la théorie analytique et probabiliste des nombres. Société Mathématique de France (1995)